Regulatory elements controlling *CisFrp1/5* expression in *Ciona intestinalis* during embryogenesis

Thesis

UNRESTRICTED

Regulatory elements controlling *CisFrp1/5* expression in *Ciona intestinalis* during embryogenesis

Palmira D'Ambrosio
Laurea in Scienze Biologiche
Università degli Studi di Napoli "Federico II"
Italia

Doctor of Philosophy

Sponsoring Establishment
Stazione Zoologica "Anton Dohrn"
Napoli, Italia

May 2006

DATE of SUBMISSION: 22 SEPTEMBER 2006
DATE of AWARD: 20 DECEMBER 2006

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Director of Studies: **Dr. Antonietta Spagnuolo**
Laboratory of Biochemistry and Molecular Biology
Stazione Zoologica "Anton Dohrn"
Naples, Italy

Second Supervisor: **Dr. Euan R. Brown**
Laboratory of Neurobiology
Stazione Zoologica "Anton Dohrn"
Naples, Italy

External Supervisor: **Dr. Stefano Piccolo**
Dept. Medical Biotechnologies
University of Padua, Italy

Table of contents

Summary

In the ascidian, endodermal progenitor cells have the capacity to differentiate autonomously from the earliest stages of embryogenesis. Their fate is restricted at the 64/110 cell stage when, after 5-6 cell divisions, the progeny of five pairs of vegetal blastomeres give rise to all endodermal cells (almost 500). The endodermal cells in the larval stage are present in the antero-ventral region of the trunk and in a strand lying on the ventral side of the tail. *Ci-titf1*, a gene homologous to mammalian *Titif1*, is a marker for early endoderm specification, during *Ciona intestinalis* embryogenesis, and for endostyle differentiation during metamorphosis.

Several studies are currently focused on markers for early specification of organs of endodermal lineage, such as liver, pancreas, pharynx and digestive tube, given their physiological importance.

My project has focused on the identification of putative *Ci-titf1* downstream genes that could represent useful markers for early regionalization of endodermal tissues in *Ciona* embryogenesis. For this purpose, through a subtractive hybridization screen between *Cititf1* overexpressing embryos and control embryos, I isolated and characterized a cDNA clone that, by sequence similarity analysis, appears to code for a factor belonging to the secreted frizzled related protein family (sFRp) and that has been called *CisFrp1/5*. sFRp proteins are modulators of the Wnt pathway, a genetic cascade that influences such diverse biological processes as developmental fate, cell polarity, adhesion, tumorigenesis and apoptosis. The expression data of *CisFrp1/5*, in the anterior region of the embryo (included the endoderm) from the neurula stage, indicate a putative involvement of this gene in late endoderm differentiation. In order to analyse the regulatory elements controlling *CisFrp1/5* tissue specific expression during

development, I used the electroporation technique to introduce reporter constructs, containing progressively deleted fragments of *CisFrp1/5* 5' promoter region fused to the *LacZ* reporter gene into fertilised *Ciona* eggs.

This strategy led to the identification of a key fragment in the promoter sequence of the gene, essential for *CisFrp1/5* activation in the areas where it is expressed. I gradually narrowed the extent of my investigation up to the identification of a 130 bp (130bp *CisFrp1/5*) element necessary for *CisFrp1/5* endoderm activation. In addition, *Cititf1* overexpression and underexpression experiments indicated that *Cititf1* could influence the transactivation of the reporter gene downstream from 130p *CisFrp1/5*, suggesting a correlation between *Cititf1* and the expression of *CisFrp1/5*.

Introduction

General

In recent years, the elucidation of mechanism by which animals from insects to higher vertebrates develop, has led to the discovery of gene products that are structurally conserved and perform functionally similar tasks. Although flies, frogs and mice are animals of obviously very different appearance, the generation of differentiated cell lineages, and of morphological and functional compartments in these organisms is governed by similar hierarchical processes. One early step in embryogenesis is the establishment of cell diversity, originating from variations in subsets of expressed genes. This seems to be controlled at least in part by the graded distribution of maternal factors. Generally, two types of action serve to translate these gradients of concentration into morphologically distinct entities. The first is the interaction of cell surface receptors with their respective ligands, which are present either on neighbouring cells or as secreted molecules. Receptor activation then induces the second step that consists of the propagation of the intracellular signal through the induction of specific transcription factors. This leads to the execution of different developmental programs within individual cells. The target genes of a specific transcription factor may in turn code for other receptors or ligands, capable of activating similar signalling pathways during later stages of development.

1 Wg/Wnt gene family

The *Wnt* family of signalling proteins participates in multiple developmental events during embryogenesis and has also been implicated in adult tissue homeostasis.

Wnt signals are pleiotropic, with effects that include mitogenic stimulation, cell fate specification, proliferation, differentiation, polarity, gene expression. Many Wnts are essential during embryogenesis, but Wnts are also active in regeneration and maintenance of adult tissue such as lymphoid tissues, colon skin, hair follicle and bone (Bienz and Clavers, 2000; Alonso and Fuchs, 2003; Staal and Clevers, 2003). Perturbations in Wnt signalling promote both human degenerative disease and cancer (Moon, 2004)..

Research on *Wnt* genes started when studies, performed by Nusse and Varmus (1982), led to a discovery of a locus termed *int*-1 (for integration site), that is activated in response to proviral insertion of mouse mammary tumor virus (MMTV). Sequence analysis showed that *int*-1 was orthologous to the *Drosophila melanogaster* segment polarity gene *wingless* (*wg*) (Baker, 1987; Cabrera *et al.*, 1987, Rijsewijk *et al.*, 1987) and the terms were combined to produce the name "*Wnt*" (the Wg-type Int-1) gene family (Nusse *et al.*, 1991). The Wnts are a family of secreted glycoproteins characterised by several conserved cysteine residues that have been identified in organisms ranging from *hydra* to humans. In vertebrates, 19 *Wnt* genes have been identified in the human genome, 16 in *Xenopus laevis*, 11 in chick, and 12 in zebrafish; in invertebrates, *Drosophila* has 7 *Wnt* genes, *Caenorhabditis elegans* 5 and *hydra* at least one (Miller, 2001). 10 *Wnt* genes are present in *Ciona intestinalis* genome (Hino *et al.*, 2003).

1.1 Wnt signalling pathways

Secreted Wnt ligands transduce paracrine signals through membrane-bound receptors, the Frizzled (Fz) proteins. Frizzled proteins (Bhanot *et al.*, 1996, Wang *et al.*, 1996) resemble G-protein-coupled receptors, with their serpentine structure, seven

transmembrane helical domains and cytoplasmic COOH-terminal. Additionally, they possess a cysteine-rich domain (CRD) at the amino terminus which is necessary and sufficient for binding to *Wnt* molecules. Wnt signalling requires also the presence of a long single pass transmembrane molecule of the LRP (Low Density Lipoproteins (LDL) receptor related protein) class, identified as the gene "arrow" in *Drosophila* (Wehrli *et al.*, 2000) and as LRP5 or LRP6 in vertebrate (Pinson *et al.*, 2000; Tamai *et al.*, 2000).

Historically, *Wnt* proteins have been grouped into two classes canonical and non canonical on the basis of their activity in cell lines or *in vivo* assays (**fig. 1**).

In the canonical pathway, binding of a particular Wnt ligand (e.g. Wnt1, Wnt3A and Wnt8) to its target Fz receptor involves formation of a membrane-located complex with one of the low-density lipoprotein receptor-related proteins LRP5 or LRP6. This interaction appears to be an early decisive event, which specifies routing of the signal through a complex regulatory transduction system. This network has been the subject of genetic, biochemical and developmental studies, especially in *Drosophila* and *Xenopus*, which led to the identification of β-catenin as a key downstream effector of Wnt/Wingless pathway. There are two pools of β-catenin in the cells: a stable membrane-associated pool involved in cell adhesion, through its interactions with cell surface proteins called cadherins, and an unstable, cytoplasmic pool. In the absence of Wnt signalling, the cytoplasmic pool of β-catenin is efficiently degraded by the activity of the serine/threonine kinase glycogen syntase kinase 3 (GSK3) (Rubinfeld *et al.*, 1996; Aberle *et al.*, 1997; Liu *et al.*, 2002). Upon binding of Wnt to Fz in complex with LRP5/6, phosphorylation and degradation of β-catenin is prevented (Yamamoto *et al.*, 1999), permitting the accumulation of the non-phosphorylated form. Stabilized β-catenin translocates to the nucleus, binds TCF/LEF factors and activates a panoply of target genes, of which more than 30 have now been identified, including c-myc, c-jun and cyclin D (Nusse, 2001). The specification of the subset of genes which is

upregulated may be influenced by the specific Wnt, Fz and other cellular and molecular contexts, and will in turn dictate the biological output of the initial signal event. During development, such outputs are critical for patterning, for example, of embryonic axis and for the formation of posterior structures (Moon *et al.*, 1997).

Noncanonical Wnts (e.g. Wnt4, Wnt5A and Wnt11) activate other signalling pathways, such as the planar-cell-polarity (PCP)-like pathway that guides cell movements during gastrulation (Heisenberg *et al.*, 2000) and the Wnt/Ca^{2+} pathway (discovered in zebrafish and *Xenopus*) (reviewed in Kuhl *et al.*, 2000) that has roles in ventralization of *Xenopus* embryos and in the regulation of convergent extension (Winklbauer *et al.*, 2001; Niehrs, 2001).

Noncanonical Wnts can even antagonise the canonical pathway (Torres *et al.*, 1996; Ishitani *et al.*, 2003) for example in the control of convergent extension (Kuhl *et al.*, 2001). However, several Wnt proteins appear to have both canonical and noncanonical properties, for example, Wnt5A, a noncanonical Wnt, induces secondary axis formation when co-expressed with its receptor, Fz5 (He *et al.*, 1997). Thus, the functional classification of Wnts may depend on the repertoire of Wnt receptors in a particular cell type.

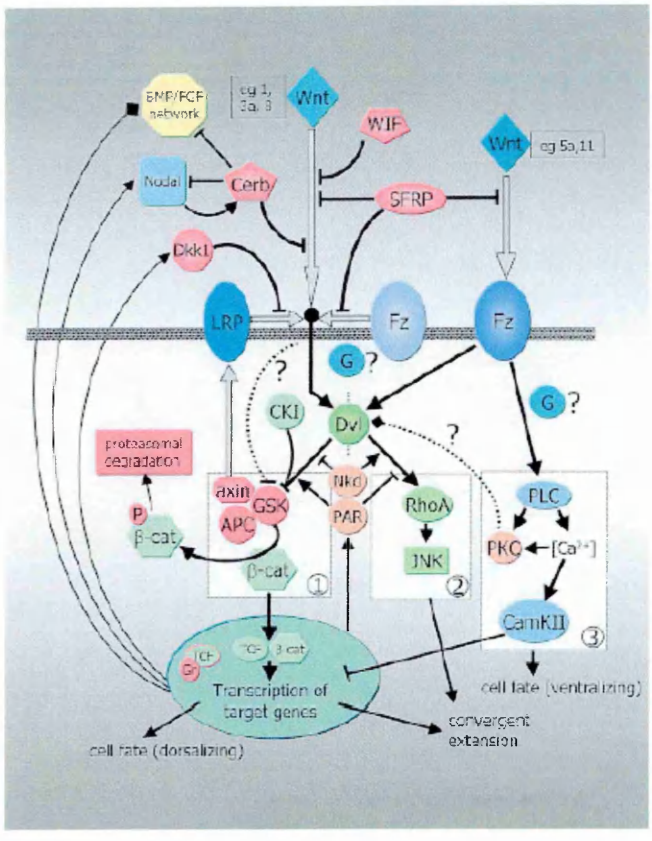


Figure 1

Schematic, partial diagram of aspects of the Wnt signalling networks (from Jones and Jomary, 2002)

Secreted, extracellular Wnt ligands signal through membrane-bound Frizzled (Fz) receptors to three major pathways in vertebrates. In the Canonical β-catenin (**Pathway 1**), certain Wnts interact with specific Fzs in complex with Lipoprotein receptor-related protein (LRP) 5 or 6 to activate the pivotal protein Dishevelled (Dvl). Dvl inhibits phosphorylation of cytoplasmic β-catenin by a complex including glycogen synthase kinase 3β (GSK), adenomatous polyposis coli protein (APC) and axin. β-catenin is translocated to the nucleus, where it relieves inhibition of transcriptional factors TCF/LEF by repressors such as Groucho (Gr). This enables increased expression of a wide range of target genes influencing cell fate and convergent extension movements. In the Polarity/Convergent extension (**Pathway 2**), different Wnt-Fz combinations signal through an alternative Dvl-mediated process in the absence of LRP and via the small G-protein RhoA, to activate c-Jun NH2-terminal kinase (JNK). Through a cascade of further signalling events, alterations in cell morphogenetic movements are regulated. In the Wnt/Ca^{2+} (**Pathway 3**), Wnt signals through Fz, again in the absence of LRP, to activate phospholipase Cβ (PLC) and lead to raised intracellular Ca^{2+} levels and activation of protein kinase C (PKC). In turn, Ca^{2+}-dependent calmodulin kinase II (CamKII) is activated and influences cell fate and cell adhesion processes.

Negative regulators (molecules and processes) of Wnt signalling are indicated in red. Secreted frizzled-related proteins (SFRP), Dickkopf-1 (Dkk1), Wnt-inhibitory factor-1 (WIF) and Cerberus (Cerb) all act extracellularly. Binding interactions are indicated by open arrows. Some possible feedback loops, potentiation paths and cross-talk with other networks are also indicated.

1.2 Secreted modulators of Wnt signalling

Since Wnt factors are very potent morphogens, they must be controlled tightly and precisely to guarantee that they function correctly both temporally and spatially. For this reason, different secreted factors are used as antagonists or "modulators" of Wnt signalling.

These molecules can be divided into two functional classes, the sFRP class and the Dickkopf class. Members of the sFRP class, which includes the sFRP family, WIF-1 and Cerberus, bind directly to Wnts, thereby altering their ability to interact with the Wnt receptor complex. Members of the Dickkopf class, which comprises certain Dickkopf family proteins, inhibit Wnt signalling by binding to the LRP5/LRP6 component of the Wnt receptor complex. Thus, in theory, those antagonists of the sFRP class will inhibit both canonical and noncanonical pathways, whereas those of the Dickkopf class specifically inhibit the canonical pathway (Fedi *et al.*, 1999; Brott and Sokol, 2002).

The first member of the sFRPs to be discovered was sFRP3/FrzB, isolated, almost simultaneously, as a chondrogenic factor in cartilage (Hoang *et al.*, 1996) and in a screen for *Xenopus* dorsal-specific factors, a screen that also led to the identification of Cerberus (Bouwmeester *et al.*, 1996). FrzB contains a characteristic cysteine-rich domain (CRD) that shares homology with the Fz CRD (**fig. 2**), which led to the prediction that it regulates Wnt signalling, as demonstrated later in *Xenopus* embryos (Leyns *et al.*, 1997; Wang *et al.*, 1997). Following these initial discoveries, further closely related secreted molecules carrying Fz-like CRD homologies were rapidly identified using a variety of approaches. In a search for EST (expressed sequence tag) databases using Fz sequences as queries, four sFRPs human genes were found (Rattner *et al.*, 1997); the corresponding murine and human cDNAs were subsequently cloned, characterized and termed sFRP1–sFRP4 (Shirozu *et al.*, 1996). From human embryonic

lung fibroblast culture a Fz gene was isolated that turned out to be identical to SFRP1 (Finch *et al.*, 1997).

In separate studies, focused on human apoptosis regulation genes, Melkonyan *et al.* (1997) identified three candidates, called SARPs (secreted apoptosis-related proteins). SARP1 (=SARP2) revealed to correspond to the already known sFRP1, sFRP2. SARP3 proved identical to a new family member, sFRP5, isolated also by Hu, *et al.*, (1998) and by Chang, *et al.*, (1999) in a search for genes expressed in pancreas or in retinal pigment epithelium of the eye. The sFRPs family has then been enriched by *Sizzled* genes (szl; for secreted frizzled), in an expression cloning approach to identify proteins that perturbed *Xenopus* embryo development (Salic *et al.*, 1997; Bradley *et al.*, 2000), and by Crescent, characterized in *Xenopus* (Pera and De Robertis, 2000) and in chick (Pfeffer *et al.*, 1997) as a gene particularly expressed in the anterior region of the embryo.

To the different sFRPs family members, discovered by so various approaches and in different model systems, were initially given various names (Jones and Jomary, 2002). Then a unifying nomenclature has been recently adopted that classify sFRPs in eight groups: sFRP1-5, found in different vertebrates, Sizzled, Sizzled2 and Crescent isolated from *Xenopus*, chick, axolotl but not identified in mammals. The generally accepted nomenclature of FRPs and their alternative names and selected properties are summarized in **Table 1**. Note that, with one exception (Illies *et al.*, 2002), sFRPs (and the other Wnt antagonists) have not been found in invertebrates.

Approved gene name	Alternative names	Species	Chromosome location		Accession nos. (human except where indicated)	mRNA sizes	Expression patterns	Protein chain length predicted MW
			Human	Mouse				
SFRP1	sFRP-1, Frp-1, SARP2, FrzA	b, h, m, r, ch, x	8p11.22	8	NM_003012	4.4 kb (h) ~5 kb (h)	Heart and most other tissues exc lung, liver	313 aa (h) 35.3 kDa
SFRP2	SDF-5, sFRP-2, SARP1	h, m, ch, rab	4q32.1 4q31.3 (Chang et al.)	3	XM_050626	~2.2 kb (b) 2.2 ÷ 1.3 kb (h)	Colon, sm. intestine, retina, heart, brain, pancreas	295 aa (h) 33.5 kDa
FRZB	sFRP-3, FrzB, Frzb-1, Fritz, Frezzled	b, h, m, ham, ch, x	2q32.1	2	NM_001463	2.3 kb (ch)	Cartilage, heart, brain, spleen, skel. muscle, kidney	325aa (h) 36.2 kDa
SFRP4	DDC-4, sFRP-4, frpAP, frpHE, FrzB-2	h, m, r	7p14.1	13	NM_003014	Not reported	Ovary, brain, kidney, lung, endometrium	346 aa (h) 39.9 kDa
SFRP5	SARP3, hFRP-1b/ Frzb-1b	b, h, m	10q24.2 10q24.1 (Chang et al.)	?	NM_003015	~1.9 kb (b) 2.2 kb (h)	Pancreas, retinal pigment epithelium	317 aa (h) 35.6 kDa
Sizzled	—	x, axl	—		[AF059570] (x)	?	Ventrolateral mesoderm	281 aa (x) 31.8 kDa
Sizzled2	—	x	—		Not found	1.2 kb (x)	Ventral embryo	280 aa (x) 31.7 kDa
Crescent	frzb2 (x)	ch, x	—		[AF136183] (x)	5.2 + 1.5 kb (x)	Organizer, dorsal embryo (x), pronephros	295 aa (x) 33.5 kDa

Table 1

Alternative names and selected properties of the SFRP genes and encoded proteins. axl, axolotl; b, bovine; ch, chick; h, human; ham, hamster; m, mouse; r, rat; rab, rabbit; x, *Xenopus* (from Jones and Jomary, 2002).

1.3 Structure/function relationships of sFRPs

sFRP proteins consist of approximately 300 amino acids, including the cysteine-rich domain or CRD, which lie in the N-terminal half of the protein, that is typically 30-50% identical to the CRDs (FRZ domain) of Frizzled receptors family members. Structurally they resemble the Frizzled receptors but lack the latter's transmembrane domains (**fig. 2**).

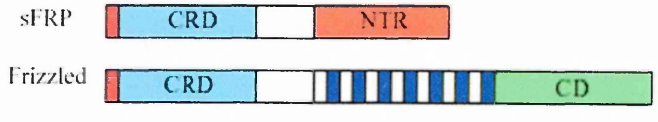


Figure 2

sFRP family proteins structure

sFRP family proteins are related to Frizzled receptors in the CRD (Cysteine-rich domain). NTR, netrin-like domain; CD, cytoplasmic domain. Signal peptides, transmembrane domains are shown as red and blue, respectively.

As demonstrated by Bafico *et al.* (1999) CRD of some sFRPs appears to interact with itself and with Fz receptor so, the sFRPs are antagonist of Wnt signalling, but it is still not known whether this effect is accomplished either by a direct interaction with Wnt proteins, to prevent them from binding to Fz receptor, or by a formation of non functional complexes with Fz. Furthermore, in the case of a direct interaction with Wnt protein, it remains unclear also which domain of sFRPs is involved, namely the CRD (Lin *et al.*, 1997) or the C terminal domain, which lies outside the CRD, (Uren *et al.*, 2000). The conflicting data result from differential affinities among sFRPs and their Wnt partners or the use of different ligands (Wg by Uren *et al.*, and Wnt-1 by Lin *et al.*).

When considering sFRP function during development, there are some examples in the literature suggesting that sFRPs could potentiate Wg activity rather than inhibiting it; these data come from tissue culture experiments using sFRP1 and Wg, the *Drosophila* Wnt homologue (Uren *et al.*, 2000). One possibility is that sFRP1 contains low and high affinity binding sites for Wg, responsible for so diverse responses. However, it must be considered that, although there is a high degree of conservation between Wg and vertebrate Wnts, the sFRP1-Wg interactions *in vitro* do not reflect those between sFRP-1 and Wnts from the same species.

The C-terminal half of sFRPs contains a domain that shares weak similarity with the axon guidance protein netrin (NTR). This NTR module, whose function is still not clear, has also been found in tissue inhibitors of metalloproteases (TIMPs) and some complement proteins (Banyai and Patthy, 1999) where the netrin domain appears to function in the inhibitory activity of these molecules against extracellular matrix metalloproteinases (MMPs). The coupling of the Fz and netrin-like domains in the sFRP family raises the possibility that their interactions with Wnts and the Wnt signalling

network may involve matrix-stabilizing activities that could have an impact on the tissue-restructuring events occurring during development.

1.4 sFRP expression pattern

The Wnt genes play fundamental roles during patterning and development of a so diverse number of embryonic structures. Since the first isolation of a sFRP gene, it appeared very important to clarify the patterns of expression of these "Wnt modulators" and to study their physiological roles. Data on the localization of sFRPs in chick, mouse, human and *Xenopus* (Jaspard *et al.*, 2000; Terry *et al.*, 2000; Esteve *et al.*, 2000; Kim *et al.*, 2001) indicate that these molecules are widely expressed throughout the development, both during tissue differentiation and early embryogenesis.

During tissue differentiation, the fields of sFRPs expression are wide and diverse.

In mouse, for example, sFRP2 signal has been detected during development of the eye, brain, neural tube, craniofacial mesenchyme, joints, testis, pancreas and below the epithelia of oesophagus, aorta and ureter, where smooth muscles develop. Moreover, in a comparative analysis, transcripts of the related sFRP1 and sFRP4 genes were frequently found in the same tissues as sFRP2 with their expression domains overlapping in some instances, but mutually exclusive in others. While sFRP1 is specifically expressed in the embryonic metanephros, eye, brain, teeth, salivary gland and small intestine, there is only weak expression of sFRP4 except for the developing teeth, eye and salivary gland (Leimeister *et al.*, 1998).

In chicken, cfrzb-1 expression at late stage of development includes neural crest cells and the condensing mesenchyme of the bones in both the limb and the trunk (Baranski *et al.*, 2000), while Sfrp-2 seems to be associated with myogenesis and Frzb with chondrogenesis (Ladher *et al.*, 2000).

In human hFRP-1b is exclusively expressed in pancreatic tissue while high levels of hFRP-2 are found in adipose tissue. In addition, low levels of hFRP-2 are also observed in other tissues including heart, pancreas and muscle. Remarkably, FRP-2 is predominantly expressed in un-differentiated preadipocytes in both rodent and man (Hu *et al.*, 1998).

As regards the early embryogenesis, the data collected so far indicate a clear involvement of some sFRPs family members in nervous system development.

In the chick, for example, where they were first analyzed, whole-mount *in situ* hybridization has revealed that the earliest expression of cfrzb-1 is in cells fated to become neural ectoderm in streak-stage embryos. Expression of cfrzb-1 in the neural ectoderm continues up through stage 8. After stage 8, cfrzb-1 expression is gradually attenuated in the closing neural tube of the trunk and is concomitantly up-regulated in neural crest cells. Finally, cfrzb-1 appears in the condensing mesenchyme of the bones in both the limb and the trunk in stage 25+ embryos (Baranski *et al.*, 2000). Sfrp-2 and Frzb are also expressed in overlapping areas in the neural plate and neural tube early in embryogenesis, while later they become localized in distinct areas, outside the nervous system (Ladher *et al.* 2000). The sFRPs territories however, are not limited to the "neural areas" in fact expression of both *sfrp1* and *sfrp2* is observed in mesodermal and ectodermal derivatives, while *sfrp1* is also found in endodermal lineages (Terry *et al.*, 2000). In the chick also the expression of *crescent* is localized in the anterior endoderm during gastrulation (Pfeffer *et al.*, 1997).

In the early mouse embryo mFrzb-1 is expressed in the brain, but also in the accessory territories including the primitive streak and the presomitic mesoderm (Hoang *et al.*, 1998).

In *Xenopus* embryos *crescent/frzb2* is expressed in the prechordal mesendoderm that, as in mouse, is strictly required for anteroposterior neural induction (Kiecher and Niehrs, 2001).

Many preliminary observations indicate that, in some cases, the patterns of sFRP expression are **complementary** to those of specific Wnts, which might support the idea that they antagonise Wnt function by delimiting the area of Wnt activity. For example, in the mouse, *sfrp1* and its potential ligand mWnt-8 exhibit identical temporal and spatial patterns of gene expression within the developing heart, and their interaction may be involved in controlling heart morphogenesis (Jaspard *et al.*, 2000).

Moreover, in the chick *Crescent* is expressed in the anterior endoderm during gastrulation when this tissue displays heart-inducing activity (Schultheiss *et al.*, 1995; Pfeffer *et al.*, 1997). At the same stage, cells in the primitive streak and posterior mesoderm, from which heart develops, express both Wnt-3a and Wnt-8c with complementary expression domains. Administration of exogenous *crescent* to posterior lateral plate mesoderm, a region of the embryo that normally forms blood island derivates, induces heart muscle formation while repressing erythropoiesis. Conversely, ectopic expression of either Wnt-8c or Wnt-3a in precardiac mesoderm blocks cardiogenesis in this tissue while promoting formation of primitive erythrocytes. These results suggest that Wnt signals present in nascent mesoderm must be blocked by Wnt antagonists secreted by anterior endoderm to permit the development of heart muscle in the anterior lateral plate of the embryo (Dale *et al.*, 1992; Jones *et al.*, 1992; Fainsod *et al.*, 1994; Graff *et al.*, 1994; Maeno *et al.* 1994; Suzuki *et al.*, 1994; Andree′ *et al.*, 1998; Schlange *et al.*, 2000).

Besides being complementary, there are some examples in the literature showing that sFRPs and Wnts can also be expressed in overlapping domains. One possibility is that sFRPs function to modulate the effect of Wnt by selectively inhibiting certain Wnts

in a localized region. SFRP-2, for example, has been shown to inhibit XWNT-8 activity (Ladher *et al.*, 2000), and SFRP-1 has been shown to selectively inhibit WNT-1 signalling but not WNT-5a signalling (Dennis *et al.*, 1999). In this model, Wnt signalling in a region may be determined by the combination of graded differences in Wnt inhibition by SFRPs and graded expression of Wnts. Indeed, it has been suggested that sFRPs act as counter-Wnts, facilitating boundary definitions, in the developing organism, of Wnt activity. A particular example has been observed in the mouse somatic mesoderm. Dorsoventral polarity of the somatic mesoderm is presumed to be established through competitive signals originating from the ventrally located notochord (Noggin and sonic hedgehog SHH-N) (Fan and Tessier-Lavigne, 1994; Johnson *et al.*, 1994; Fan *et al.*, 1995; McMahon *et al.*, 1998) to specify the sclerotome, while dermomyotome appears to receive input from an array of Wnt proteins expressed in the dorsally located surface ectoderm (Wnt4 and Wnt6) and dorsal neural tube (Wnt1 and Wnt3a) (Fan *et al.*, 1997; Capdevila *et al.*, 1998; Marcelle *et al.*, 1997). The dorsal polarity is presumably maintained by the balance activities of the two signals, which likely display opposing concentration gradient across the somatic field (Fan *et al.*, 1997). sFRP2 is the only member expressed in the sclerotome and it is up regulated and utilized by SHH-N to exclude Wnts activity in the ventral somites.

It is also possible that sFRPs at times function as carrier proteins for Wnts for long-range diffusion or as protector proteins to prolong the half-life of Wnts until they encounter their receptors.

Another possibility is that the overlapping patterns of expression of sFRPs simply reflect the regulation of sFRP expression by Wnts. *Sfrp2* expression in the aggregating mesenchyme, for example, is induced by Wnt-4, which is critical for kidney development at this early stage (Lescher *et al.*, 1998). The potential competitive interaction by direct binding of sFRP to Wnts may concentrate Wnts within localized

region. This could potentiate Wnt activity by increasing its local concentration in a given area.

The outcome of all these studies is that there are contradictory data and unresolved issues, mostly related to the recent discovery of this gene family, to the limited functional approaches and the limited number of model systems used to study their roles. As already mentioned, with the exception of the sea urchin *Strongylocentrotus purpuratus* (Illies *et al.*, 2002), it seemed that sFRPs (and the other Wnt antagonists) were not present in invertebrates.

Recently, the sequencing of the genome of the ascidian *Ciona intestinalis* (Dehal *et al.*, 2002), that is my model system, revealed the presence of four sFRP genes in this chordate ancestor, whose study could provide important clues for understanding their function and their regulation during embryogenesis. The identified four sFRP genes are *Cis*FRP1/5, assigned to the first group including mammalian sFRP1 and sFRP5, *Cis*FRP2, assigned to the second group including mammalian sFRP2, CisFRP3/4-b and *Cis*FRP3/4-a assigned to the third group including mammalian sFRP3, sFRP4 and *Xenopus* Frzb.

2 The experimental model

Ascidians (sea squirts) or tunicates are sessile marine invertebrate chordate that belong to the most primitive branch of the chordate phylum, the Urochordate, which diverged from the last common ancestor of all chordate at least 520 million years ago (Wada and Satoh, 1994; Cameron *et al.*, 2000) (**fig. 3**). This divergence creates over one billion years of independent evolution between extant ascidians and modern vertebrates, such as human, mouse, chick, frog, and zebrafish. Despite this evolutionary distance, the basic features of the chordate body plan remain recognizable in ascidian larvae. The

ascidian *Ciona intestinalis* has occupied centre stage in evolutionary and developmental biology. This is because the ascidian tadpole larva is believed to represent the closest living form to the ancestral chordate; therefore, discovering the molecular mechanisms that underlie the development of this larva would facilitate the understanding of the ancestor of humans and other chordates.

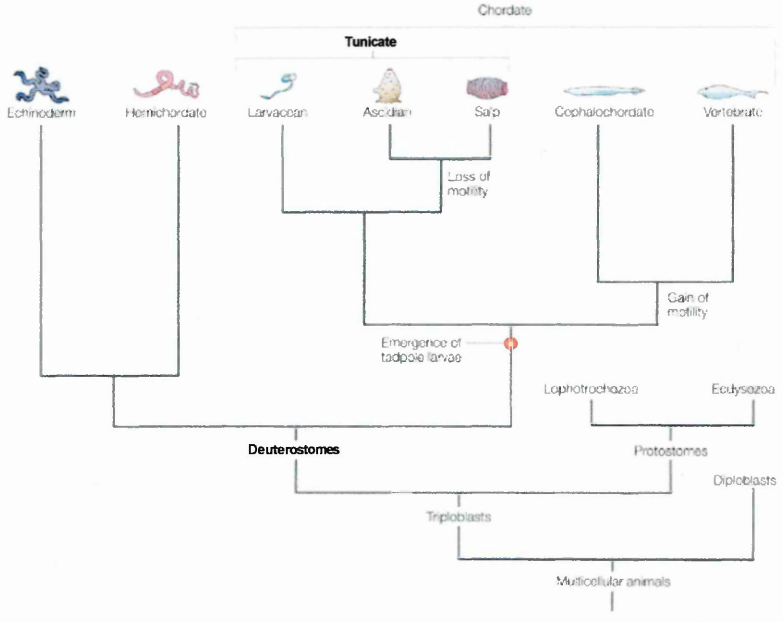


Figure 3

The evolution of chordate

The chordates comprise the tunicates, the cephalocordates and the vertebrates. They are thought to have evolved from a common ancestor shared with the non-chordate deuterostomes (the echinoderms and the hemicordates). The emergence of tadpole-type larvae was a key event in the evolution of chordates. From Satoh (2003).

The close relationship between ascidians and vertebrates was first recognized in the mid-19th century by the embryologist Alexander Kowalevsky, who noted the striking similarities between ascidian larvae and vertebrate embryos (Kowalevsky, 1866). Notably, the tail of the ascidian tadpole contains a prominent notochord and a dorsal neural tube. These findings provided evidence that ascidians are, along with vertebrates and the cephalochordate Amphioxus, members of the phylum Chordata. The adult

ascidian also possesses recognizable chordate features. It is a sedentary filter-feeder, and the feeding basket of the adult contains gill slits that appear to share a common origin with the gill slit of other chordates (Aros and Viragh, 1969). Likewise, the endostyle of the adult ascidian is a homolog of the vertebrate thyroid gland.

2.1 Why ascidians?

In addition to their peculiar evolutionary position at the transition between the non-chordate deuterostomes and the chordates, the ascidians provide a simple experimental system for the investigation of the molecular mechanisms that underlie cell-fate specification during development for several reasons. First, the ascidian tadpole consists of only ~2,600 cells, which constitute a small number of tissues including the epidermis, the central nervous system (CNS) with two sensory organs (otolith and ocellus), the endoderm, mesenchyme, trunk lateral cells (TLCs) and trunk ventral cells (TVCs) in the trunk, and the notochord and muscle in the tail (**fig. 4**).

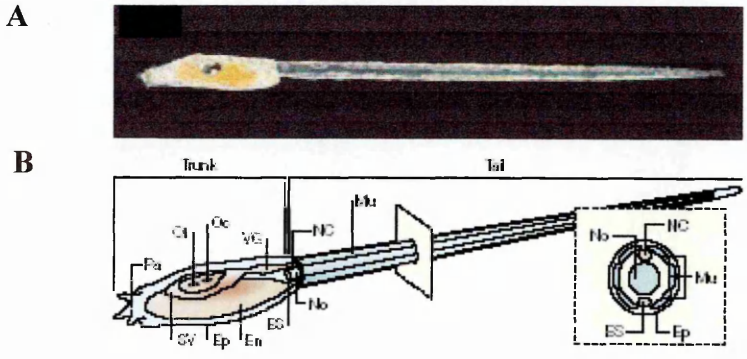


Figure 4

***Ciona intestinalis* tadpole larva**

Picture (**A**) and schematic representation (**B**) of the larval anatomical characteristics. En, endoderm; Ep, epidermis; ES, endodermal strand; Mu, muscle; NC, nerve cord; No, notochord; Oc, ocellus; Ot, otolith; Pa, palps; SV, sensory vesicle; VG, visceral ganglion (from Satoh N., 2003).

Embryogenesis of ascidians is simple. The cleavage pattern is invariant, and cleavage is bilaterally symmetrical. Gastrulation is initiated around 118-cell stage, and it involves epibolic movements of ectodermal cells and migration of endodermal and mesodermal cells inside the embryo. Neurulation is accomplished by folding of the presumptive neural plate, as in vertebrate embryos (Satoh, 1978). In addition the lineage of the embryonic cells is well documented and characterized by detailed description of the epidermis, CNS, endoderm, mesenchyme, TLCs, muscle, and notochord (Conklin, 1905; Ortolani, 1955; Nishida and Satoh, 1983; Nishida, 1987; Nicol and Meinertzhagen, 1988). During ascidian embryogenesis the developmental fate is restricted early, between 64/110 cell stage, when each blastomere gives rise to a single specific type of tissue at the larval stage. Cloning and characterization of developmental genes indicate that each gene is expressed under discrete spatio-temporal pattern within their lineage (Yasuo and Satoh, 1993; Satou *et al.*, 1995; Di Gregorio and Levine, 1998; Satou and Satoh, 1999; Wada and Satoh, 2001).

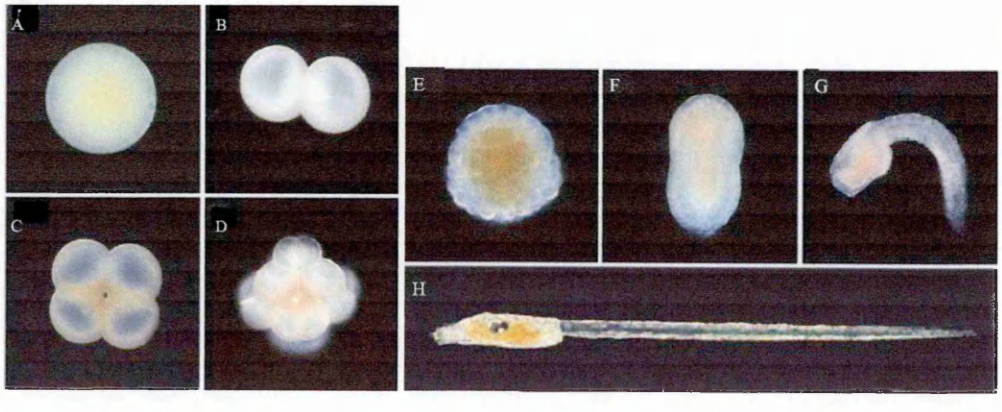


Figure 5

Stages of *C. intestinalis* embryogenesis

Fertilized egg (**A**), 2-cell embryo (**B**), 4-cell embryo (**C**), 16-cell embryo (**D**), gastrula (**E**), early-tailbud embryo (**F**), mid-tailbud embryo (**G**) and tadpole larva (**H**). Embryos were dechorionated to show their outer morphology. From Satoh N., 2003

Furthermore, the blastomeres of early ascidian embryos are large, easy to manipulate and allow the detailed visualization of changes in gene expression during development. Additionally, *Ciona* embryogenesis is rapid (taking ~18 hours from fertilization to the development of a free-swimming tadpole, at 18°C) (**fig. 5**) and *C. intestinalis* develop into mature reproductive adults by 3 months, facilitating genetic analyses. Novel functions of developmental genes can be determined by misexpressing or overexpressing a variety of regulatory genes that encode transcription factors and signalling molecules, or by the functional suppression of genes with morpholino oligonucleotides. Finally, transgenic DNA can be introduced into developing ascidian embryos using simple electroporation methods. This strategy is considerably more efficient than conventional microinjection assays and permits the simultaneous transformation of hundreds, of synchronously developing embryos.

2.2 Ascidian genomics

The unique evolutionary position of the ascidians - as invertebrate chordate - makes their genome particularly interesting from an evolutionary point of view. The release of the draft sequence of the *C. intestinalis* genome (Dehal *et al.*, 2002) showed that its 153-159 Mb genome (20 times smaller than the human genome) contains approximately 16.000 protein-coding genes, a number that is similar to that predicted for other invertebrates and only half of that predicted for vertebrate genomes. An annotated assembly of the genome is freely available through the JGI's Web site (http://genome.jgi-psf.org/ciona/), providing a resource that greatly accelerates the identification of homologs of genes previously studied in other organisms. In addition to the sequence of the genome, a large-scale expressed sequence tag (EST) project has been carried out, resulting in the characterization of about 18.000 independent cDNA

clones, estimated to represent about 85% of the *C. intestinalis* transcripts (Satou *et al.*, 2002). The results of this EST survey are available through the Kyoto University Web site (http://ghost.zool.kyoto-u.ac.jp/indexr1.html). The small size of the *C. intestinalis* genome provides a distinct advantage for understanding genome organization and gene function. It has been proposed that large-scale gene duplications occurred in the vertebrate lineage after it had diverged from the cephalochordates and the urochordates. Ascidians have a basic non-duplicated chordate type genome, with most of the minimal promoters located <1 kb upstream of the transcription start sites of the genes. This feature of the ascidian genome, together with the electroporation method mentioned previously, has made them particularly useful in studies on the function and on the transcriptional regulation of genes controlling embryogenesis and germ layer specification, a key question in developmental biology.

Of the three primary germ layers, the ectoderm and mesoderm have traditionally received considerably more experimental attention than the endoderm (Smith, 1989; Kimelman *et al*, 1992; Kessler and Melton, 1994; Slack, 1994; Heasman, 1997). Recently, a confluence of studies on different model systems, included ascidians, has served to correct this historic imbalance and to elucidate fundamental molecular mechanisms underlying endoderm development.

3 Endoderm: definition and general considerations

Endoderm is one of the three primary germ layers formed in the embryo; it moves inward by cell movements during gastrulation. The term gastrulation means "formation of the gastric cavity", the archenteron, a first hint of a digestive structure, accomplished through the definitive differentiation of groups of blastomeres in endodermal cells. In vertebrates, as the embryo folds, the definitive endoderm forms a

tube within the developing body. This tube will undergo a set of morphogenetic changes that lead to formation of various organs involved in digestion, respiration, hormonal balance and waste disposal. Endoderm differentiation therefore represents a fundamental step during embryo morphogenesis and one can legitimately suppose that this event is temporally linked to gastrulation. Up to gastrulation the endoderm remains anatomically linked to the mesoderm, through anatomic proximity and shared molecular mechanisms. The insights, from studies in several model systems, have led to the concept that the early events of differentiation are characterised from an initial separation of a "mesoendodermal field" (Kimelman and Griffin, 2000) from the ectoderm. In a second step the mesendoderm segregates into endoderm or mesoderm depending on the appropriate combination of signalling in each compartment. The origin of the vertebrate endoderm has been delineated in the early amphibian embryo, where pigmented cells of the animal pole form the ectoderm while yolky cells of the vegetal pole develop into endoderm. Transplantation and culture experiments of disaggregated blastomeres in *Xenopus* embryo have demonstrated that when a single vegetal pole cell from either the morula (stage 6) or midblastula is isolated and transplanted into the blastocoel of the late blastula, its progeny is found in all three germ layers (Heasman *et al.*, 1984). When vegetal pole cells from early gastrula are assayed in this way, however, their progeny contributes only to gut endoderm (Heasman *et al.*, 1984). These results suggest that vegetal pole cells become committed by the beginning of gastrulation and further analysis of their progenies in more detail have shown that this commitment is gradual (Wylie *et al.*, 1987).

Zebrafish fate map studies began in the 1990s and these fate maps showed that both endoderm and mesoderm also originate from a common progenitor. Both germ layers derive from cells near the blastoderm margin (Kimmel *et al.*, 1990) and both involute into the forming hypoblast (Warga and Kimmel, 1990). At the onset of

gastrulation, the majority of the endoderm progenitor cells are the earliest deep involuting cells from the blastoderm margin in the newly formed hypoblast (Warga and Nüsslein-Volhard, 1999). The fate commitment of the endoderm also occurs just after the onset of gastrulation (Ho and Kimmel, 1993; David and Rosa, 2001). When marginal cells from late blastulae (30–40% epiboly) are transplanted into animal blastomeres, they contribute mostly to neuroectodermal tissue, consistent with the fate maps of animal blastomeres, and only a small proportion of transplanted cells contribute to the endoderm (David and Rosa 2001). This contribution to the endoderm dramatically increases, however, when marginal cells from embryos are transplanted at the onset of the gastrulation (50% epiboly) (David and Rosa, 2001).

Mammalian embryos produce two forms of endoderm: primitive (visceral), which colonizes extraembryonic tissues, and definitive, which contributes exclusively to fetal tissues (Gardner, 1982). The most definitive endoderm cells originate from the anterior part of primitive streak that is necessary for gastrulation to occur properly and also seems to be involved in cell fate specification of mesendoderm progenitor cells. During gastrulation, the mesoendodermal cells accumulate at the anterior end of the primitive streak, involute through the primitive streak, migrate along the midline and give rise to mesoderm and definitive endoderm. The endoderm precursor cells are thought to intercalate into overlying visceral endoderm layer, eventually displacing the cells from the visceral layer.

The origin of endoderm and mesoderm from the same precursor cells is conserved across metazoan phylogeny.

In the nematode *Caenorhabditis elegans*, the endodermal germ layer is established as a single cell, called E, which derives from the division of the mesoendodermal precursor cell (EMS). The EMS divides into an MS cell (that produces

mesodermal muscles) and an E cell (that produces intestinal endoderm) (Thorpe *et al.*, 1997).

A mesoendodermal field can be also identified in sea urchin, where both endoderm and mesoderm derive from the lower part of vegetal pole of the embryo (Ruffins and Ettensohn, 1996).

Experimental evidence from studies in these animal models suggests a close correlation between gastrulation and endoderm specification and that the cell commitment is a gradual process. In an initial phase, after separation of mesendoderm from ectoderm, maternal factors within the mesoendodermal field promote the endoderm and mesoderm differentiation. In a second phase, downstream effectors generate and stabilize a programme for commitment to endoderm cell lineage.

3.1 Endoderm formation in vertebrate: molecular mechanisms

Recently, a number of powerful studies have begun to characterize the molecular determinants of the endoderm, a germ layer previously neglected in developmental biology. Work in diverse vertebrate model systems, as *Xenopus*, zebrafish and mouse (Stainier, 2002), were done concurrently, profoundly influenced one another and has converged on an integrated transcriptional and signalling pathway that serves to establish the vertebrate endoderm.

The data accumulated so far, indicate that the process starts by the action of a T-box transcription factor, VegT, identified so far only in *Xenopus* (Zhang *et al.*, 1998; Clements *et al.*, 1999).

VegT mRNA is largely localized in future endoderm area in the vegetal hemisphere of the egg and early embryo (Lustig *et al.*, 1996; Stennard *et al.*, 1996; Zhang and King, 1996; Horb and Thomsen, 1997), and it induces endoderm when

expressed ectopically in animal cap explants (Horb and Thomsen, 1997). Selective depletion of the maternal store of *Xenopus VegT* mRNA abrogates differentiation of all endoderm and most mesoderm, defects that are corrected by re-expressing VegT (Zhang *et al.*, 1998).

VegT, in *Xenopus*, exerts its function by activating and reinforcing the expression of *Nodal*-related TGF signalling molecules (*TGFβ/Nodal*: *Xnr1*, *Xnr2*, *Xnr4*, *Xnr5*, *Xnr 6* and *Derriere*) (Kofron *et al.*, 1999; Hyde and Old, 2000, Takahashi *et al.*, 2000). The correlation between VegT and *TGFβ/Nodal* came from experiments that tested the ability of various *TGFβ/Nodal* family members, such as *Xnr1*, *Xnr2*, and *Xnr4* as well as *derrière*, to restore endodermal gene expression in VegT-depleted embryos (Xanthos *et al.*, 2001).

Homeobox containing genes of the Mix/Bix family (*Mixer*, *Mix.1*, *Mix.2*, *Bix1*, *Bix2/Milk*, *Bix3 e Bix4*) and Gata family members are the downstream effectors of VegT and VegT/Nodal (Weber *et al.*, 2000; Afouda *et al.*, 2005).

Furthermore, these proteins cooperate for regulating the expression of a relay of HMG-box Sox-family transcription factors culminating with *Sox 17*, which may be an obligate mediator of vertebrate endoderm development (Xanthos *et al*, 2001; Zhang *et al.*, 2005).

Genetic studies in zebrafish have confirmed the data obtained in *Xenopus*. A Nodal mutant, *OEP* (*one-eyed-pinhead*), lacks endoderm (Schier *et al.*, 1997; Zhang *et al.*, 1998), as a double mutant for two other nodal factors, *Cyclops (cyc)* e *Squint (sqt)* (Feldeman *et al.*, 1998).

In mouse, the first *Nodal* gene was identified through the study of a retroviral insertion that, when homozygosed, leads to the lack of a primitive streak and most mesendoderm as well as to an early arrest of embryonic development (Iannaccone *et al.*, 1992; Conlon *et al.*, 1994).

As already mentioned, downstream effectors of Nodal gene are homeobox Gata and Sox transcription factors.

A confirmation of the involvement of Gata factors in endoderm formation has come from the analysis of the zebrafish *faust* (*fau*) mutation that encodes zebrafish Gata5 and is required for endoderm formation upstream of *Sox17* (Reiter *et al.*, 1999, 2001). In addition, *fau/gata5* is sufficient to induce *Sox17* expression and requires the function of the novel Sox protein Cas (Kikuchi *et al.*, 2001; Reiter *et al.*, 2001). Despite *gata4*, *gata5*, or *gata6* widespread and early expression in this tissue, genetic analyses in mouse have so far failed to reveal a clear role for these genes in endoderm formation (Molkentin, 2000). It is expected that tissue specific gene inactivation and/or the simultaneous inactivation of several of these genes may reveal their likely role in this process. **Figure 6** illustrates the molecular regulatory cascade leading to endoderm induction in zebrafish and *Xenopus*.

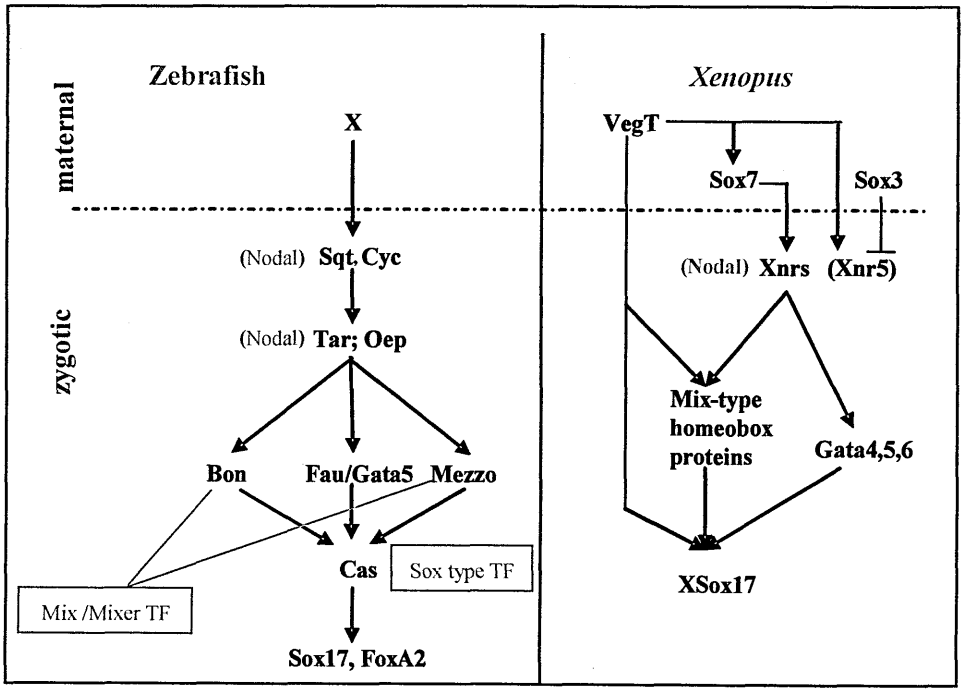


Figure 6

Molecular pathway leading to endoderm in the zebrafish and *Xenopus*.

Maternal factor (X) upstream of Nodal (*Cyc, Sqt*) is unknown. Squint is a maternal transcript. Tar; Tarama, type-I TGFβ receptor (Renucci *et al.*, 1996). Xnrs; Xnr1, Xnr2, Xnr4, Xnr5 and Xnr6. Mix-type homeobox proteins; Mixer, Mix.1, Mix.2, Bix.1, Bix.2, Bix.3 and Bix.4.

Members of the *Mix/Bix* and *Sox17* families, furthermore, have been cloned in *Zebrafish* and in mouse (Hudson *et al.*, 1997; Henry and Melton, 1998; Dickmeis *et al.*, 2001; Kikuchi *et al.*, 2001; Kanai *et al.*, 1996) and their molecular function seem to be conserved across the phyla. In *Zebrafish*, for example, *bon* (a Mix gene) acts upstream of *Sox17* (Alexander and Staineir 1999); *bon* mutants, indeed, lack *Sox17* and 90% of endodermal tissues (Kikuchi *et al.*, 2001). *Sox17* mutant in mouse embryos shows profound deficits in gut (Kanai-Azuma *et al.*, 2002).

The studies conducted so far point, therefore, to conserved molecular pathways across vertebrate phyla, although the hierarchical relations amongst the different factors need to be clarified, given their functional redundancy.

3.2 Endoderm formation in invertebrates: molecular mechanisms

Most of the current knowledge on endoderm formation in lower organisms came from experiments in *C. elegans*, *Drosophila* and, more recently, in ascidians.
While these studies have revealed a high degree of conservation between vertebrates and invertebrates, as to some of the transcriptional regulators, they also have highlighted the divergence in the intracellular signalling events leading to formation of this tissue. Transcription factors as GATA are a clear example of a gene family implicated in endoderm formation across the phyla.

In *Drosophyla*, some of the factors involved in endoderm differentiation have been isolated through mutations that affect the formation of this tissue. They include *huckebein*, *serpent*, and *fork head. huckebein* encodes a Zn-finger protein (Bronner *et al.*, 1994). It interacts with *snail* and *twist* to separate the endoderm from the mesoderm prior to gastrulation. At the molecular level, Huckebein has been proposed to repress the function of Snail and Twist in the endodermal progenitors (Reuter and Leptin, 1994).

Moreover, Huckebein regulates positively the expression of the *GATA* gene *serpent* in the midgut primordia and *serpent* itself is required for midgut development (Reuter 1994). *fork head* mutants also lack a gut and *fork head* encodes the founding member of the winged-helix transcription factors, now known as Fox (Forkhead box) factors (Weigel *et al.*, 1989; Kaestner *et al.*, 2000).

In *C. elegans* the entire endoderm originates from the E (Endoderm) blastomere at the 8-cell stage (**fig. 7**) (Sulston *et al.*, 1983). E itself is a daughter of EMS and the sister of MS (Mesoderm), which gives rise to much of the mesoderm. Endoderm formation in *C. elegans* is under the control of a number of maternal proteins, including SKN-1, a bZIP/homeodomain transcription factor that is required for both E and MS development (Bowerman *et al.*, 1992). SKN-1 directly regulates the expression of two redundant GATA-type transcription factor genes, *med-1* and *med-2*, which like

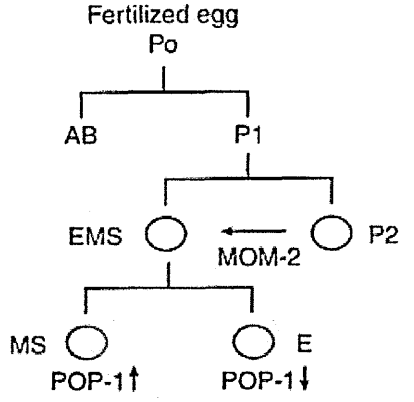


Figure 7

Early cell lineage of the *Caenorhabditis elegans* embryo

In *C. elegans*, the mesendodermal progenitor EMS divides to give rise to MS, the mesodermal progenitor, and E, the endodermal progenitor. Prior to this division, P2 sends EMS a Wnt signal (encoded by *mom-2*) which leads to a decrease in POP-1 function in E, allowing endoderm formation.

skn-1 are necessary for EMS differentiation (Maduro *et al.*, 2001). END-1 and END-3, another pair of redundant GATA factors, lie downstream of MED-1 and MED-2 (Zhu *et al.*, 1997). Shortly after the time E is originated and *end-1* and *end-3* become activated,

two additional GATA factors, ELT-2 and ELT-7 become expressed (Fukushige *et al.*, 1998).

Upstream from all these GATA factors however a Wnt signal appears to be fundamental for the "choice" between endoderm and mesoderm lineages in *C. elegans*. MOM2 (a Wnt homolog), secreted by P2 cell, interacts with MOM5 receptor (a Frizzled homolog) on the EMS cell (Thorpe *et al.*, 1997); the signal blocks the expression of POP1 (a TCF homolog) in the E cell (which is closest to the P2 cell) while in the MS cell, that does not receive the Wnt signal, the nuclear level of POP1 is high (Lin *et al.*, 1998).

The Wnt signalling and the GATA transcription factors have been implicated in mesendoderm differentiation also in the sea urchin. The mesendoderm specification starts when β-catenin, responding to an unknown signal, accumulates in the nuclei of all mesendodermal precursors and forms a complex with TCF (Logan, 1999). This complex regulates the expression of the signalling molecule Wnt8, which in turn activates the β-catenin/TCF signalling creating a positive feedback loop (Angerer e Angerer, 2000). The β-catenin/TCF complex also activates several transcription factors including krox, otx and krl, a kruppel like factor as early response genes (Howard *et al.*, 2001; Wang *et al.*, 1996) and later endodermal genes, as GataE, foxA and fox.B At the early blastula stage GataE becomes the major endodermal regulator, as demonstrated by morpholino antisense experiments (Davidson *et al.*, 2002).

Therefore, the formation of a primitive tissue like the endoderm seems to conserve most of the transcriptional regulators across the phyla and to diverge in the signalling molecules used for establishing the expression domains of these transcriptional regulators. Signalling molecule as Tgfβ/Nodal appear to be vertebrate specific, while the signalling activated by Wnt molecules seems to be specific for endoderm differentiation in invertebrates.

Indeed, in vertebrates, several lines of evidence demonstrate that the Wnt/β-catenin pathway plays an essential role upstream of endoderm specification: β-catenin localization in the vegetal pole of the embryo, in fact, is the first event for dorsal axis formation and for mesendoderm separation from the ectoderm in *Xenopus*, zebrafish and mouse (Schneider *et al.*, 1996; Schohl and Fagotto 2002; Beddington and Robertson, 1999; Lu *et al.*, 2001).

3.3 Endoderm specification in Ascidians

In ascidians, the canonical Wnt pathway, via the regulation of β-catenin, is essential for the differentiation of the endoderm, retaining the ancestral role as in worms (Imai *et al.*, 2000). The endoderm of an ascidian larva is a tissue present mostly in the trunk region and formed by almost 500 cells, whose lineage has been completely described; Satoh, 1994). Since the first experiments by Ortolani (1971), largely confirmed by Nishida (1993), the endoderm differentiation appears as an autonomous process, as demonstrated by the potential of presumptive endodermal blastomeres to differentiate autonomously when isolated from the early embryo (Nishida, 1992; Wittaker, 1990). This potential reflects the presence of still uncharacterized, maternal pre-localized cytoplasmic factors, present in the endoplasm of eggs and early embryos (Nishida 1993). Their distribution has been investigated by cytoplasmic transfer experiments involving the fusion of blastomeres and non-nucleated egg fragments (Nishida, 1993, 1994; Yamada and Nishida, 1996). The importance of these factors has been recently confirmed, in *Halocynthia roretzi*, by Kobayashi and Nishida (2001) who demonstrated, by egg cytoplasm transfer, that blastomeres committed to muscle fate can be induced to express the alkaline phosphatase, an unequivocal endodermal marker in ascidians (Whittaker, 1990). This occurs even after the recipients have already restricted

to a specific cell type and after the expression of the tissue-specific gene is already initiated. Lineage studies of endodermal blastomeres have been conducted by intracellular injection of the tracer enzyme horseradish peroxidase (HRP), a well established technique for tracing cell lineages in ascidians (Nishida, 1987). The maternal endodermal determinants are inherited by the two pairs of vegetal blastomeres A4.1 and B4.1 in the 8 cells embryo. After three cell divisions, at the 64 cells stage, the endodermal fate is restricted to 5 blastomere pairs that, together with a pair of blastomeres, fate restricted at the 110 cells stage, after 5-6 cell divisions give rise to the approximately 500 endodermal cells of the larva. Recent studies have suggested a pivotal role for endodermal tissue in inducing differentiation of both notochord and mesenchyme cells. Experiments in *H. roretzi*, involving coisolation and recombination of blastomeres at different developmental stages, have shown that cellular interactions with adjacent presumptive endodermal blastomeres, during the late 32-early 64 cells stages, are required for notochord and mesenchyme formation (Kim and Nishida, 1999; Kim *et al.*, 2000). The inductive molecules from endodermal cells, identified as a FGF factor in *H. roretzi* (Nakatani *et al.*, 1996) and a Notch (Corbo *et al.*, 1997a, 1998) or a FGF factor (Imai *et al.*, 2002) in *C. intestinalis*, act synergistically with *macho-1* (a muscle determinant) for mesenchyme development and with still unknown factors to induce notochord differentiation (Imai *et al.*, 2002; Kobayashi *et al.*, 2003). Without these signals, notochord precursors adopt a neural fate, while mesechyme precursors give rise to muscle cells (Kobayashi *et al.*, 2003). Maternal determinants, therefore, supply the blastomeres with "basic differentiation programs" that can be altered by external signals. Furthermore, studies on *Ciona* and *H. roretzi* have suggested that the blastomeres of endoderm and notochord lineages have the potential to differentiate in both direction and that the final differentiation is induced by the appearance of either *Brachyury* (for notochord) or *Cititf1* (for endoderm) (Yasuo and Satoh, 1998;

Ristoratore *et al.*, 1999). This indicates that also in ascidians, as in higher chordates and in lower invertebrates, it may be envisaged the existence of a mesendodermal field that separates early from the ectoderm, and that later gives rise to the endoderm and mesoderm, depending on the appropriate combination of specific signals and factors in each territory. At the molecular level, the accumulation of β-catenin in the nuclei of vegetal blastomeres at the 32-cell stage, is the first step, identified so far both in *C. intestinalis* and in *C. savignyi*, in the process of ascidian endoderm specification (Imai *et al.*, 2000). Mis-and/or over expression of β-catenin induces the production of an endoderm-specific alkaline phosphatase (AP) in presumptive notochord cells and epidermis cells, without affecting differentiation of primary lineage muscle cells. Downregulation of β-catenin, induced by the overexpression of cadherin (the adhesion molecule able to sequestrate the cytoplasmic pool of free β-catenin), results in the suppression of endoderm cell differentiation (Imai *et al.*, 2000). A subtractive hybridization screen between β-catenin- overexpressed embryos and cadherin-overexpressed embryos has led to the identification of potential β-catenin target genes involved in endoderm differentiation in *Ciona savignyi* (Satou *et al.*, 2001; Imai, 2003). The downstream genes so identified have been classified into three groups. One group represents genes that are involved in endoderm formation (*FoxA5*, *Lhx3* and *Titf1*) the second group includes genes that are relevant for embryonic induction of mesodermal tissues (*Fgf9/16/20*, *chordin*, *FoxD*, *Zic*, *Brachyury*, *twist-like1*, *Mesp*). The third group contains genes whose functions have not been fully analyzed in ascidians (*cadherinII, protochadherin*, *Eph*, *lefty*, *dkk*, *DMRT1*, *hairy*, *ELK*, *Fli, jun*, *msxb*, etc.).

As to the first group, both *Lhx3* and *Titf1* have been deeply characterized, while the role of *FoxA5* (former *forkhead/HNF3*β), that is broadly expressed in vegetal blastomeres as early as the 16-cell stage (*Mocu-FH1*, Olsen and Jeffery, 1997; *Hr-HNF3*, Shimauchi *et al.*, 1997; *Cifkh*, Di Gregorio *et al.*, 2001), remains to be further elucidated.

At the 32-cell stage, the expression of *Cs-Lhx3*, a LIM-class homeobox gene starts (Wada et al., 1995). Initially the signal is not strictly restricted to endoderm blastomeres, while it becomes localized into endoderm blastomeres by the 110-cell stage (**fig. 8**). (Satou *et al.*, 2001).

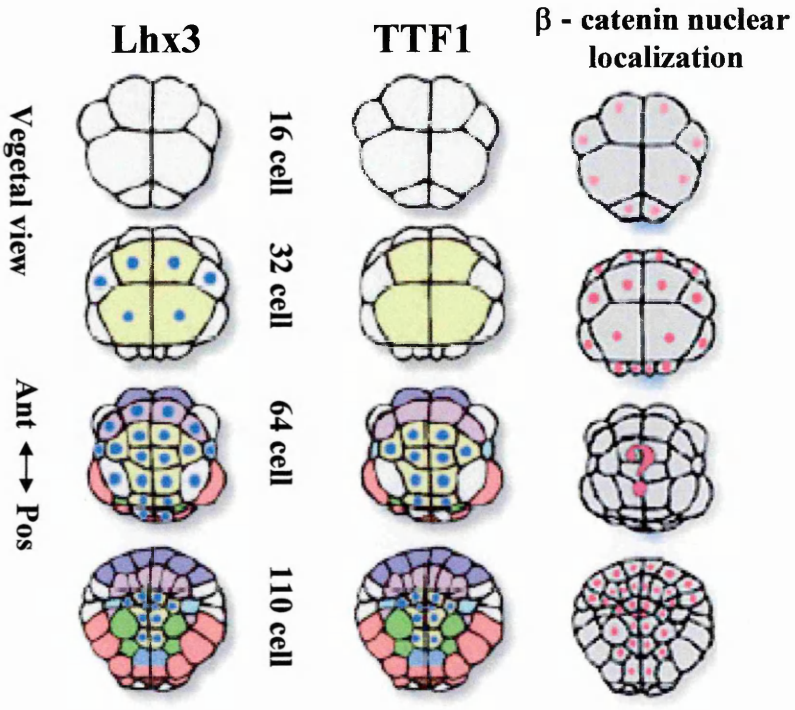


Figure 8

Zygotic gene expression of Lhx3 and TTF1 genes during cleavage stages.
Expression of each gene is indicated by blue dots on blastomeres. On the right nuclear localization of β-catenin during cleavages is shown. Expression of the gene is indicated by red dots on blastomeres.

Cs-Lhx3 is required and sufficient for the expression of late marker genes of endoderm differentiation. Overexpression of *Cs-Lhx3* promotes ectopic endoderm formation without β-catenin activity; furthermore, *Cs-Lhx3* can restore alkaline phosphatase expression in β-catenin depleted embryos (Satou *et al.*, 2001; Nishida, 2005).

The first zygotic factor detected exclusively in endodermal cells since the 64 cell stage, is *Cittf1*, a transcription factor containing an NK-2-like homeodomain (Ristoratore *et al.*, 1999). Its expression, localized in the endodermal territories up to the gastrula stage, disappears at the neurulae stage, and reappears in endodermal precursor cells at the tailbud and larval stages; after metamorphosis *Cittf1* is detected in the endostyle. The mRNA injection of *Cittf1* causes ectopic expression of alkaline phosphatase in blastomeres of notochord lineage and results in the development of larvae with abnormalities in trunk-tail development consequent to the formation of excess endoderm (Ristoratore *et al.*, 1999). *Cititf1* ectopic expression, exclusively in notochord blastomeres, using the promoter of a notochord-specific gene, *Brachyury* (*CiBra*-promoter), results in tadpole larvae phenotypically similar to that obtained by *Cititf1* mRNA injection (Spagnuolo and Di Lauro, 2002). Furthermore, *Cs-ttf1* expression is upregulated in the β-catenin mRNA-injected 110-cell embryos and is suppressed in the *Cicadherin* mRNA injected embryos at the 110-cell stage (Satou *et al.*, 2001), indicating *Cs-ttf1* as a downstream effector of β-catenin cascade in endoderm differentiation.

Hence *Cititf1* not only represents the first homeobox containing gene of the *NK-2* class, shown to directly affect on endoderm development, but it is also the first specific regulatory endodermal marker to be isolated from an ascidian. Moreover, *Cititf1* seems to play additional important roles later during endoderm patterning and endostyle differentiation (Ristoratore *et al.*, 1999).

Embryos of ascidians develop into swimming tadpole larvae that metamorphose into sessile adult forms. *Ciona* larvae, as other ascidian species, do not feed since they are not provided of functional digestive organs; trunk endoderm appears as a mass of undifferentiated cells rich in yolk granules. Although no sign of organogenesis is evident, it has been demonstrated that, after metamorphosis, developmental fates of

larval endodermal cells, are almost fixed, indicating that a certain degree of regionalization, in this tissue, already exists (Hirano and Nishida, 2000). *Cititf1*, a marker for early endoderm specification, marks also endostyle differentiation during metamorphosis: at the larval stage its mRNA is, indeed, present in the anterior-ventral part of head endoderm, a region that, after metamorphosis, give rise to the endostyle of the adult (the thyroid ancestor) where *Cititf1* is still present. *Cititf1*, therefore, is a clear confirmation of a regular projection from the cleavage to the larval stage and from the larval to the adult stage.

Aim

Cititf1 is the first specific endodermal marker isolated from the ascidian *Ciona intestinalis* (Ristoratore *et al.*, 1999). Previous studies demonstrated that ectopic expression of *Cititf1* in the notochord, using *Brachyury* promoter (*CiBra/Cititf1* transgene), results in the development of larvae with abnormalities in trunk-tail development consequent to the formation of excess endoderm (Spagnuolo and Di Lauro, 2002) (**fig. 9**).

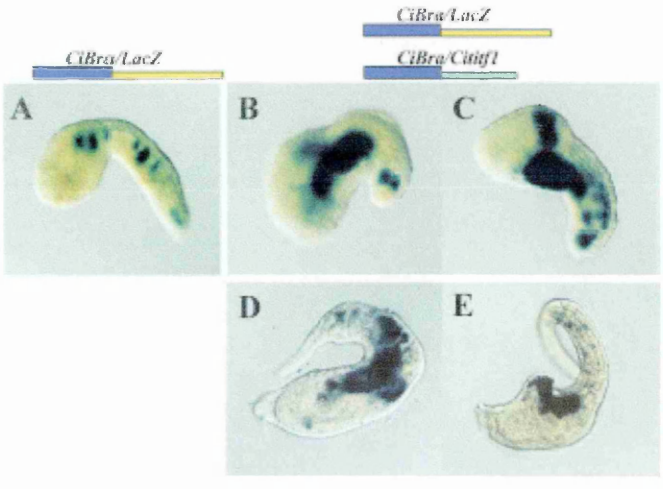


Figure 9

Misexpression of *Cititf1*.

The *CiBra* promoter region was attached upstream from *Cititf1* cDNA (*CiBra/Cititf1*), to express *Cititf1* ectopically in the notochord. As control *CiBra/LacZ* was used.

A) Control embryo, electroporated with *CiBra/LacZ* at tailbud stage shows LacZ staining localized specifically in the notochord. The embryos in (**B**,**C**), which were instead coelectroporated with the same reporter gene plus the *CiBra/Cititf1* transgene, show an altered LacZ staining pattern and a mutant phenotype. The aberrant development become more evident at larval stage (**D**, **E**) where the embryos show an altered trunk development and a shorter, bent tail missing an organized notochord structure.

These phenotypically altered embryos were used as starting material for a subtractive hybridization screen to identify genes that could be target of *Cititf1* and eventually involved in early regionalization of endodermal territories from neurula to tailbud stages. Amongst different cDNA clones surveyed, I focused my attention on a secreted molecule, *CisFrp1/5*, whose expression is restricted in the anterior region of the embryos from the neurula stage. sFrps are proteins, recently discovered in vertebrates, which are to be involved in Wnt pathway. Most of their functions are still undiscovered and only scanty information on the mechanism controlling their regulation is available. Moreover, with the exception of the sea urchin *Strongylocentrotus purpuratus*, *sFrps* are considered to be absent in invertebrates.

CisFrp1/5 therefore represented a starting point to start to study these genes in an invertebrate chordate, the *C. intestinalis*. For this purpose, firstly I defined its expression pattern during *Ciona* embryogenesis; then I focused my attention on *CisFrp1/5* transcriptional regulation. I identified a region of about 1600 bases able to drive the expression of the reporter gene in the territories where the endogenous gene is expressed. This region has been restricted, through a series of deletion constructs, to a fragment of about 200 bases. Moreover I tried to find a correlation between *Cititf1* and *CisFrp1/5* activation, through *Cititf1* overexpression and underexpression experiments.

Methods

Ascidian eggs and embryos

C. intestinalis adults were taken in the bay of Naples. For *in vitro* fertilization, eggs and sperm were collected from the gonoducts of several animals. Fertilized eggs and embryos were used in electroporation or *in situ* hybridisation experiments. Embryos were raised in Millipore-filtered seawater at 18–20 °C. Only the batches in which 90% or more of the embryos developed normally were selected for the experiments. Samples at appropriate stages of development were also collected by low-speed centrifugation and used for RNA extraction.

RNA isolation

Total RNA was extracted from embryos with Eurozol (Euroclone) following the manufacturer's protocols. 140µl of Eurozol was used for about 700 embryos. After chloroform-isoamilic extraction and isopropanol precipitation, the RNA pellet was air-dried and re-suspended in 15 µl of diethyl pyrocarbonate-treated (DEPC) water. After the addition of 1µl SUPERaseIn (Ambion 20U/µl), 2µl of transcription buffer (Ambion 10X) and 2 µl of DNAseI RNAse free (10U/µl) the samples were placed in a heating block at 37°C for 30 minutes. After this incubation 30µl of DEPC was add to the reaction solution, water and total RNA was extracted with 55µl phenol pH 4.3, 15µl chloroform-isoamilic alcohol and 5µl NaOAc 2M pH 4. After centrifugation at 10,000 rpm for 10 minutes at 4°C in Eppendorf microcentrifuge, the top (aqueous) layer was taken out and placed in a clean tube, after the addition of the same volume of choloroform: isoamylalcohol (24:1) it was vortexed thoroughly and centrifuged at

13,000 rpm for 10 minutes. Then the aqueous layer was taken out and placed into a new tube. Precipitation was carried out with the addition of 1 volume isopropanol and 0.5 µl glicogen. The RNA pellet was air-dried and re-suspended in 20 µl of diethyl pyrocarbonate-treated (DEPC) water and quantified by spectrophotometer measurement. An aliquot of each sample was conserved for gel electrophoresis and for inspection of RNA quality (i.e. 28S:18S RNA ratio).

First-Strand cDNA synthesis

About 1 µg total RNA was used to synthesize the first strand of cDNA with SMART (Switching Mechanism at 5' end of RNA Transcript) polymerase chain reaction (PCR) cDNA Synthesis Kit (Clontech Company). The SMART PCR cDNA synthesis technology utilizes a combination of two modified primers in a single reaction. The modified oligo dT primer (CDS primer: 5'-AAGCAGTGGTAACAACGCAGAGTACT(30)N-1N-3') is used to prime the first-strand reaction, while the SMART oligonucleotide (SMART primer : 5'-AAGCA-GTGGTAACAACGCAGAGTACGCGGG-3') serves as a short, extended template at the 5' end of the RNA template. The resulting full-length, single-strand (ss) cDNA contains the complete 5' end of the mRNA and the sequence complementary to the SMART oligonucleotide, which then serves as a long distance PCR priming site to amplify the full-length cDNA. The first-strand of cDNA was used to generate the double-stranded cDNA by long distance PCR with Advantage 2 Polymerase Mix (CLONTECH Laboratories, Inc). Total RNA was reverse transcribed to first strand cDNA with PowerScript Reverse Transcriptase. Long Distance (LD) PCR of cDNA amplification was optimized depending on the amount of total RNA used in the first strand synthesis and thermal cycler. To determine the optimal PCR cycles with

Eppendorf Gradient thermal cycler, PCRs were performed from 15 to 30 cycles. More PCRs were done after analysis through 1.2 % agarose gel electrophoresis to determine the exact PCR cycles of every sample. Placental total RNA was performed as control. cDNA was then phenol extracted according to standard procedure, concentrated in n-butanol, purified upon chroma spin-1000 columns (Clontech) and eluted in TNE buffer (10 mM Tris-HCl pH 8.0, 10 mM NaCl, 0.1 mM EDTA).

Digestion with *Rsa*I and purification of digested products

cDNAs prepared from RNA of *Cititfl* overexpressing embryos, electroporated with gene *CiBra/Cititfl* transgene, control embryos electroporated only with *CiBra/LacZ* and control cDNA from human placental tissue were treated with *Rsa*I (New England Biolabs, Beverly, MA, USA), a four-base recognition site restriction enzyme, that was chosen because it generated the largest average size of fragments (about 600 bp). The digested cDNA were purified using Nucleo Trap (Clontech), ethanol precipitated and adjusted to 300 ng/µl in TNE buffer. To control cDNA digestion and purification, uncut and *Rsa*I-digested as well as unpurified and purified cDNAs were electrophoresed on 1.5% agarose/ethidiumbromide gels.

Generation of a subtracted library by subtractive suppression hybridization (SSH)

To enrich cDNAs up-regulated in *CiBra/Cititfl* electroporated embryos, SSH was performed between cDNA prepared from *CiBra/Cititfl* transgene embryos (tester) and cDNA prepared from control embryos (driver) using PCR-Select cDNA subtraction kit according to the protocol (Clontech, Von Stein *et al.*, 1997). In brief, 0.5 mg of the cDNA tester was ligated to 40 pmol of two single-stranded phosphorylated adaptors

(adaptor1:5'CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT 3';adaptor2R:5'CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGT 3') (ShengYou Inc., China) by T4 DNA ligase (New England Biolabs). Ligation efficiency was analysed by amplifying DNA fragments spanning the adaptor/cDNA junctions by PCR using primer-1 and glyceraldehyde3-phosphate dehydrogenase (G3PDH) primers. Primer sequences were: G3PDH 5'-primer 5'-ACC ACA gTC CATgCC ATC AC-3'; G3PDH 3'-primer 5'-TCC ACC ACC ATgTTg CTg TA-3'; primer-1 5'-CTA ATA CgA CTC ACTATA ggg C-3'. The adaptor-ligated cDNA tester was then hybridized for two rounds with excessive cDNA driver. The hybridized products were first amplified with primer 1 in a 50 μl reaction containing 2 μl of SSH products, PCR buffer (20 mM Tris/HCl pH 8.4, 50 mM KCl and 1.5 mM $MgCl_2$), 1 mM primer 1, 0.1 U/ml Taq polymerase, 200 mMdNTPs, under the following cycling conditions: filling the adaptors for 5 min at 75°C; denaturation for 2 min at 94°C; 32 cycles of 30 s at 94°C, 45 s at 66°C and 1.5 min at 72°C; a final extension for 5 min at 72°C. Aliquots of 3 ml of PCR products were then secondarily amplified in a 100 μl reaction containing PCR buffer, 1 μM nested PCR primer 1, 0.1 U/ml Taq polymerase and 200 μM dNTPs, under the following cycling conditions: 32 cycles of 33 s at 94°C, 45 s at 68°C and 1.5 min at 72°C; a final extension for 5 min at 72°C. After 18, 23, 28 and 33 cycles, 5 μl aliquots of each PCR tube were electrophoresed on a 2% agarose ethidium bromide gel. The cDNA was then directly cloned into pCR 4-TOPO vector according to the manufacturer's instructions (Invitrogen, NV Leek, The Netherlands). Briefly, 10 ng of the secondary PCR product was ligated into 10 ng vector, and the ligation mixture was introduced into competent TOP10 bacteria by electric shock. The library was plated onto LB plates containing 50 μg/ml ampicillin. Single bacterial transformants were picked randomly and grown in 96-well plates in 100 μl LB medium containing 50 μg/ml ampicillin at 37°C for at least 4 h.

Analysis of sPCR inserts

Substraction PCR (sPCR) probes were amplified using M13 forward and M13 reverse primers (35 cycles of 94°C for 1 min, 48°C for 2 min, and 72°C for 3 min). Finally, an extension cycle was carried out at 72°C 5-20 minutes to complete all the strands. The amplification cycles were conducted by means of Thermal cycler Perkin-Elmer-Cetus. The PCR products were purified using a QIAquick PCR purification kit, and samples sequenced using T3 primer.

DNA sequencing

The DNA sequences were obtained using the Beckman CEQ 2000XL DNA Analysis System Apparatus by the Molecular Biology Service at the Stazione Zoologica "A. Dohrn" in Naples.

Isolation of cDNA clones from ordinate library

The full length cDNA of subtractive cDNA clones and *CisFrp1/5* full length cDNA were isolated from an arrayed cDNA collection gently provided by Professor Noriyuki Satoh from Department of Zoology, Graduate School of Science, Kyoto University, Sakyo-Ku, Kyoto, Japan. The gene collection was prepared from different cDNA libraries obtained from RNA poly(A)$^+$ extracted from different stages, from eggs to larvae, by sequencing about 500,000 ESTs categorised into about 18,000 independent cDNA clusters that were re-arrayed in about 36 384-well plates and released from Noriyuki Satoh's laboratory. The gene collection was opportunely ordered with the aim to have each clone identified by the CloneIDs, which are two specific coordinates to

pick the cDNA clone corresponding to the gene we are interested in. The vector used for the library construction was pBlueScript SK and the inserts were cloned between *Eco*RI and *Xho*I sites with 5'-3' orientation. The clones of interest were collected from the ordered plates stored at -80°C and grown at 37°C overnight in 5 ml LB in the presence of ampicillin (50 μg/ml) in order to prepare the DNA.

Computational tools for sequence analyses

DNA sequences were compared with the sequence collection deposited in GenBank and EMBL, by means of the Genetics Computer Group (GCG) TFastA program (Devereux *et al.*, 1984; Pearson and Lipman, 1988) and were blasted against *Ciona* 1.0 database in the Doe Joint Genome Institute (JGI) *C. intestinalis* site (http://genome.jgi-psf.org/ciona4/ciona4.home.html) (Dehal *et al.*, 2002) where it is possible to find the genomic sequences, the gene model predictions, the identity degrees with the homologues of other species and the expressed sequence tags (ESTs) from the *C. intestinalis* cDNA project (http://ghost.zool.kyoto-u.ac.jp/indexr1.html) (Satou *et al.*, 2002). Genomic comparison between *C. intestinalis* (http://genome.jgi- psf.org/ciona4/ ciona4.home.html) and *C. savignyi* (http://www.broad.mit.edu/annotation/ciona/) *CisFrp1/5* gene sequences was carried out using the mVISTA server (http://www-gsd.lbl.gov /vista/mvista /submit.shtml) (Mayor *et al.*, 2000). One sequence is selected as a base or reference sequence. The server automatically uses RepeatMasker to mask repetitive elements in the reference sequence. The X-axis of the generated plot represents the base sequence and the Y-axis represents the percentage identity.

DNA digestion with restriction endonucleases

Plasmid DNA was digested with the suitable restriction endonuclease in a mixture containing 5 units enzyme/1 μg DNA, in a final volume which was at least 20 times larger than the enzyme volume in the presence of 1/10 of a suitable buffer and at a specific temperature as suggested by manufacturer's instructions.

Random Priming

For the preparation of radiolabeled DNA probes corresponding to the cDNA clones differentially expressed, dsDNA was denatured and placed in the presence of a mixture of random olgonucleotides, 6 bases in length, which act as primers for the polymerisation by *E. coli* Klenow DNA polymerase I. The reaction was performed in the presence of [α-^{32}P]dCTP and [α-^{32}P]dATP which were incorporated in the new synthesized DNA. 1.5 pmol of random primers and sterile H_2O up to 37 μl were added to 40 ng of DNA template. The template DNA was denatured at 95°C for 5 minutes. The following components up to a volume of 65 μl: 50 μM dGTP, 50 μM dTTP, 0.5 mg/ml BSA, 200 μM HEPES pH 6.6, 200 μM Tris-HCl pH 7, 10 μM EDTA, 30 μCi [α-^{32}P]dATP (3,000 Ci/mmol), 30 μCi [α-^{32}P]dCTP (3,000 Ci/mmol), 3 units of Klenow DNA polymerase I were added to the mixture at RT. The reaction was conducted at 37 °C for 2 hours. Successively, in order to separate the labelled DNA from the free precursors, the mixture was placed in Sephadex G-50 (Pharmacia) column equilibrated with 1x TE (10 mM Tris-HCl, 1 mM EDTA pH 8) and 0.1% SDS. 150 μl fractions were collected and 1 μl from each was counted by LS1701 Beckman scintillation counter, using 5 ml of Insta-Gel (Packard). The fractions containing the radioactivity peak were mixed and utilised for the hybridisation.

Hybridisation with cDNA

Hybridisation was conducted using $1x10^6$ cpm/ml of labelled probe in a mixture containing: 5x Denhardt's solution, 5x SSC, 0.5% SDS, 5 mM EDTA, 50 mM Na-phosphate pH 6.8, 100 µg/ml sonicated salmon sperm (SSS). The filter was incubated in the mixture 2 hours for the pre-hybridisation at the same temperature as for the hybridisation. Before the addition, the probe was denatured at 95 °C for 5 minutes. The hybridisation at high stringency conditions occurred at 65 °C overnight. Washes were carried out twice at the same temperature in 2x SSC and 0.1 % SDS for 15 minutes and then in 0.2x SSC and 0.1% SDS in the same conditions. The hybridisation at low stringency conditions occurred at 55°C and washes were carried out at the same temperature in 2x SSC and 0.1 % SDS. The filters were finally exposed overnight with a Kodak X-AR film at -80 °C.

Southern blot analysis

DNA was transferred on nylon filter by treating the gel 10 minutes with 10 mM HCl, 15 minutes with a denaturing solution (0.5 M NaOH, 1.5 M NaCl) and 2x15 minutes with a neutralising solution (1.5 M NaCl, 0.5 M Tris-HCl pH 7.5, 15 mM EDTA). The gel was then washed in 10x SSC 30 minutes, placed on a glass layer and a nylon filter (Hybond-N, Amersham), equilibrated in 10x SSC, was placed on it. A pile of Whatmann 3MM paper and a 250 g weight were placed on the filter. The transfer of the DNA to the filter occurred overnight at RT and DNA was permanently linked to the filter by means of U.V. exposure in Stratalinker 2400 (Stratagene). The filter was hybridised with different cDNA clones as described previously.

***In vitro* transcription**

In order to synthesise sense and antisense mRNA corresponding to the cDNA fragments isolated from the *C. intestinalis* substractive cDNA library, the plasmid containing the promoters T3 and T7 respectively upstream and downstream of the cloning sites was first linearised. When the plasmid was digested with an enzyme cutting at the 3' of the insert, RNA polymerase recognised T3 promoter and a sense mRNA was obtained. On the contrary, when an enzyme cutting at the 5' was used, the T7 promoter recognized so that an antisense mRNA was synthesised. The plasmid was digested, separated by electrophoresis and extracted from gel as previously described, then treated with 50 µg/ml protease K and 1% SDS at 45°C 30 minutes to remove possible protein contaminants. An extraction was carried out in 1 volume of phenol:chlorophorm:isoamyl alcohol (25:24:1) and a second extraction in 1 volume of chlorophorm:isoamyl alcohol (24:1). The DNA was precipitated 30 minutes at -80°C with Na-Acetate pH 5.2 up to final 0.3 M and 3 volumes cold 95% ethanol and successively centrifuged at 10,000 rpm at 4 °C 30 minutes and resuspended in DEPC H_2O. Aliquots were analysed on agarose 1 % in TBE 1x.

In vitro transcription was carried out by using a DIG-RNA labeling kit (Boehringer). By this method, it is possible to label mRNA with digoxygenin, a steroid compound isolated from plant *Digitalis planaria.* This molecule acts as hapten, which is covalently bound, by an arm containing 11 carbons, to the RNA synthesis precursor UTP C-5 position. The precursor DIG-UTP will be incorporated in the RNA synthesis.

Sense and antisense mRNAs were produced in the following reaction mixture: linearised plasmid DNA 1 µg, 10x NTP mix, 2 µl; 10x transcription buffer 2 µl, 20 U/µl RNase inhibitor 1 µl, 20 U/µl T7 RNA polymerase 2 µl or 20 U/µl T3 RNA polymerase 2 µl, DEPC H_2O up to 18 µl. Synthesis reaction was conducted at 37°C 2 hours, then 2

µl of 10 U/µl DNase I Rnase-free were added and incubated 10 minutes at 37°C to remove DNA template. The reaction was stopped by adding 1 µl of 0.5 M EDTA pH 8.

Riboprobe quantification

To estimate RNA concentration, an immunoassay was performed by means of alkaline phosphatase-conjugated anti-digoxygenin antibody. Once interaction between antibody and corresponding hapten (DIG-UTP) took place, it is possible to visualize RNA molecules by means of a colorimetric reaction catalysed by alkaline phosphatase. This enzyme produced an insoluble blue substrate in the presence of 2 enzymatic substrates, 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) and nitroblue tetrazolium (NBT). By comparing the intensity of the spots produced by different dilutions of a control RNA with the ones relative to the RNA sample, it is possible to determine its concentration. RNA dilutions were prepared using the dilution buffer [DEPC H_2O:20x SSC:formaldehyde (5:3:2)]. 1 µl of each dilution was placed on nylon filter then exposed 30 seconds to U.V. radiations to let RNA bind to the filter. It was washed 2 minutes in 2x SSC on slow shaking, then 30 minutes in blocking solution, the same in which the anti-DIG AP antibody (0.15 U/ml) will be added and incubated 1 hour at RT. To remove unbound antibodies, two washes in a solution containing 0.1 M maleic acid pH 7.5 and 0.15 M NaCl were done. The filter was equilibrated in the detection solution (100 mM Tris-HCl pH 9.5, 100 mM NaCl, 50 mM $MgCl_2$) and incubated in the dark in the same solution in which BCIP (50 mg/ml) and NBT (50 mg/ml) were added. The coloured compound started to precipitate in few minutes and the reaction was blocked after 10 minutes by washing the filter with H_2O.

Whole mount *In situ* hybridisation (WISH)

In situ hybridisation was carried out on *C. intestinalis* whole mount embryos at the interested stages fixed at RT 90 minutes in a mixture containing: 4% paraformaldehyde, 0.1 M MOPS pH 7.5, 0.5 M NaCl. Embryos were washed 2x10 minutes in PBS and dechorionated by hand with subtle platinum needles. The samples were washed 3 x 15 minutes in 1 ml PBT (PBS + 0.1% Tween 20) at RT and incubated 30 minutes at 37 °C in 1 ml PBT containing 4 µg/ml proteinase K to increase the permeability of the cells and accessibility to mRNA target. After digestion, samples were washed 3x15 minutes in PBT and post-fixed 1 hour at RT in 1 ml 4% paraformaldehyde in PBS. Embryos were again washed 3x5 minutes at RT in PBT, placed 10 minutes at same temperature in the hybridisation solution (50% formamide, 5x SSC, 50 µg/ml tRNA, 5x Denhardt's solution, 0.1% Tween 20, 50 µg/ml heparin) and PBT (1:1), 10 minutes in the hybridisation solution and finally 1 hour at 55 °C. Riboprobe was added to a final concentration of 0.5 ng/µl and the hybridisation occurred overnight at the same temperature. The following day a series of washes was conducted at different temperatures and salinity to avoid unspecific interactions. Embryos were washed 2x15 minutes in 4x SSC, 50% formamide, 0.1% Tween 20 and again 2x15 minutes in 2x SSC, 50% formamide, 0.1% Tween 20. Successively, embryos were treated 3x10 minutes at 37°C with 1 ml of solution A (0.5 M NaCl, 10 mM Tris-HCl pH 8.0, 5 mM EDTA, 0.1% Tween 20), treated with RNase A (20 µg/ml) 30 minutes at 37 °C, then washed again 15 minutes with 1 ml of solution A at the same temperature. Another series of washes was performed 20 minutes in 2x SSC, 50% formamide, 0.1% Tween 20, and 2x15 minutes in 1x SSC, 50% formmide, 0.1% Tween 20 at 45°C, then 15 minutes in SSC:PBT (1:1), and 4x5 minutes in PBT. Before adding the anti-DIG antibody, embryos were incubated 30 minutes at RT in PBT containing 5% sheep serum. The anti-DIG antibody (750 U/ml) is diluted 1:2000 in an identical

solution. 500 µl of this solution were added to each sample incubated at 4°C overnight on a shaking machine. The samples were washed at RT 4x10 minutes in PBT, 2x5 minutes in AP buffer (100 mM NaCl, 60 mM $MgCl_2$, 100 mM Tris-HCl pH 9.5, 0.1% Tween 20). In order to localise the DIG-labeled RNA bound to anti-DIG antibodies conjugated with alkaline phosphatase, two substrates were added which were converted by the phosphatase into a blue precipitate. The two substrates were nitro-blue-tetrazolium (NBT) and 5-bromo-4-chloro-3-indoxyl phosphate (BCIP) and, in particular, each sample was incubated in 1 ml of AP buffer containing 4.5 µl of NBT and 3.5 µl of BCIP. Timing of precipitate formation depended on the amount of bound antibodies. When a signal was detected, the colorimetric reaction was stopped in PBT.

Oligonucleotide synthesis

Synthetic oligonucleotides were prepared by a Beckman SM-DNA Synthesizer apparatus by the Molecular Biology Service at the Stazione Zoologica "A. Dohrn" in Naples.

Bacterial cell electroporation

By cell electroporation, it is possible to transform bacterial cells with plasmids containing DNA of interest. Briefly, the circular plasmid DNA and competent bacterial cells, prepared by the Molecular Biology Service of the Stazione Zoologica "A. Dohrn", were placed in an electrocuvette. The electrocuvette was subjected to an electric pulse at constant 1.7 V using a Bio-Rad Gene Pulser™ electroporation apparatus. Then the cells were placed in 1 ml of SOC (tryptone 20 g/l, yeast extract 5 g/l, 5 M NaCl 2 ml/l, 1 M KCl 2.5 ml/l, 1 M $MgSO_4$ 10 ml/l, 1 M $MgCl_2$ 10 ml/l, 1 M glucose 20 ml/l) shaking at

270 rpm at 37°C for 1 hour, plated on LB solid medium (NaCl 10 g/l, bactotryptone 10 g/l, yeast extract 5 g/l, agar 15 g/l) in the presence of the specific antibiotic (50 μg/ml) to which the plasmid is resistant, and then grown at the same temperature overnight.

DNA Mini - and Maxi-Preparation

A single bacterial colony containing the plasmid DNA of interest was grown in a suitable volume of LB in the presence of the appropriate antibiotic shaking at 37°C overnight. The QIAGEN® Plasmid Purification kit, based on alkaline lyses method, was used to isolate the plasmid DNA from the cells according to the manifacturer's instructions.

PCR screening

It is possible to carry out a PCR reaction using as template the DNA of a single bacterial colony and in the same time grow the colony. Each single colony was half placed in a PCR mixture described below, and half grown in 3 ml of LB liquid medium in the presence of the suitable antibiotic (50 μg/ml) 8-12 hours shaking at 37°C. The PCR reaction had the following composition: synthesis buffer 2 μl; 2 mM dNTP 2 μl, 20 pmol primer up, 20 pmol primer down, 5 U/μl Taq DNA polymerase 0.2 μl, H_2O up to 20 μl. After lysis at 95°C 5 minutes, 25 amplification cycles were structured as follow: denaturation at 94°C 50 seconds, annealing at 55-65°C 50 seconds, extension at 72°C 1-2 minutes. By electrophoresis analysis, the samples presenting a band of expected size were identified and plasmid DNA was purified from the corresponding bacterial colonies.

PCR amplification

Amplification reaction on *C. intestinalis* spermatozoa genomic DNA was conducted in a mixture containing: 100ng/µl DNA template, 5 µl 10x synthesis buffer, 5 µl 2 mM dNTP, 60 pmol primer up, 60 pmol primer down, 0.5 µl Taq DNA polymerase (5 U/µl final volume), H_2O up to 50 µl. The amplification cycles were conducted by means of Thermal cycler Perkin-Elmer-Cetus. After denaturation at 95°C 5 minutes, 25-35 amplification cycles were structured as follows: denaturation at 94°C 30 seconds, annealing at 55°C 35 seconds, extension at 72°C 3 minutes. Finally, an extension cycle was carried out at 72°C 10 minutes to complete all the strands.

Gel extraction

The DNA fragments were excised from the agarose gel and extacted by means of the QIAGEN's QIAquick Gel Extraction Kit according to the manufacturer's instruction.

***In silico* analysis of putative CiTTF-1 binding site**

The *CisFrp1/5* genomic sequence of 1680 bp was submitted to the TRANSFAC database (http://transfac.gbf.de/TRANSFAC) (Heinemeyer *et al.*, 1999). This is a relational database of transcription factors from many organisms, their genomic binding sites and DNA-binding profiles. By this analysis, two putative specific *CiTitif1* binding site were found.

Preparation of electroporation constructs

The vector used in the preparation of constructs was pBlueScript II KS containing the *LacZ* and SV40 polyadenylation sequences (pBSLacZ) (Locascio *et al.*, 1999) and all the genomic fragments were inserted in the 5'-3' orientation (**fig. 10**).

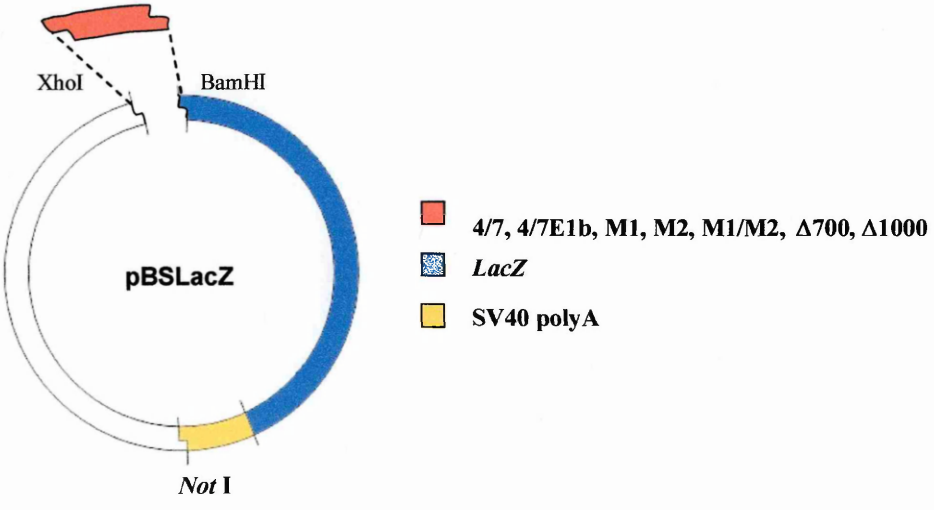


Figure 10

Cloning strategy in pBlueScript/LacZ/SV40 (pBSLacZ) vector

The vector pBlueScript, containing the coding sequence of *LacZ* (reporter gene) and the SV40 poyadenilation signal (pBSLacZ), was used to clone different *CisFrp1/5* promoter fragments (in red) in the *Xho*I and *Bam*HI restriction sites.

The genomic fragments which extends -1681 bp to -65 bp upstream the ATG starting codon, contained in the 4/7 construct (**fig. 25**), was amplified, as previously described, by PCR reaction from *C. intestinalis* spermatozoa genomic DNA. The oligonucleotides, oligonucleotide 4 and oligonucleotide 7, used for the amplification were designed according to the sequence available on JGI *C. intestinalis* site

(http://genome.jgi-psf.org/ciona4/ciona4.home.html) (Dehal *et al.*, 2002). Amplified fragments were analysed and extracted from agarose 1% gel as already described. PCR product was, then, cloned in pCR® II-TOPO®. The clones in which the insert was in 3'-5' orientation were selected by PCR screening, in order to have the *Bam*HI and *Xho*I sites, present in the pCR® II-TOPO® polylinker, in 3'-5' orientation (4/7 TOPOII construct). Successively, to prepare the construct 4/7E1b (**fig. 26**) the same genomic sequence was amplified using the same oligonucleotides but, in particular, the reverse primer, oligonucleotide 7E1b, was designed with E1BTATA sequence at 5'. Amplified fragment, analysed and extracted, was cloned in pCR® II-TOPO®. By PCR screening, described above, the positive clones were selected (4/7E1b TOPOII construct). This plasmid was used as starting vector to prepare the follow constructs. Successively, the fragments and the vector (pBSLacZ) were digested with *Xho*I and *Bam*H1, analysed and extracted from agarose 1% gel and then ligated with a 1:3 ratio vector/insert.

Mutant vectors M1, M2 and M1/M2 (**fig. 29a**) were obtained by site-directed mutagenesis following the Quick-Change Site Directed Mutagenesis Kit (Stratagene) instructions, using plasmid 4/7E1b as template and appropriate primers, each containing the desidered adjacent 3 mutated base residues as instructed by the manufacturer.

The construct Δ1000, containing the genomic fragment -981 bp upstream the ATG starting codon, (**fig. 30**) was prepared in two steps. In the first step the starting vector 4/7E1bTOPOII was digested with *Nde*I and and *Not*I, to eliminate 700 bases. Successively, the non cohesive ends were filled-in and ligated. The resulting plasmid was digested with *Xho*I and *Bam*H1. In the second step, the fragment *Xho*I -*Bam*H1, analysed and extracted from agarose 1.2% gel, was ligated to pBSLacZ, digested with the same enzymes, with a 1:3 ratio vector/insert.

The genomic fragment between -1681/-981 bp upstream the ATG starting codon, contained in the construct Δ700 (**fig. 30**) was amplified by PCR reaction from

the genomic fragment of the starting vector 4/7TOPO II. The oligonucleotides used for the amplification were designed according to the sequence of the genomic DNA and contained E1BTATA sequence at 5'. The amplified fragments were analysed and extracted from agarose 1% gel and then cloned in pCR® II-TOPO® (700/TOPO II construct). The clones in which the insert was in 3'-5' orientation were selected by

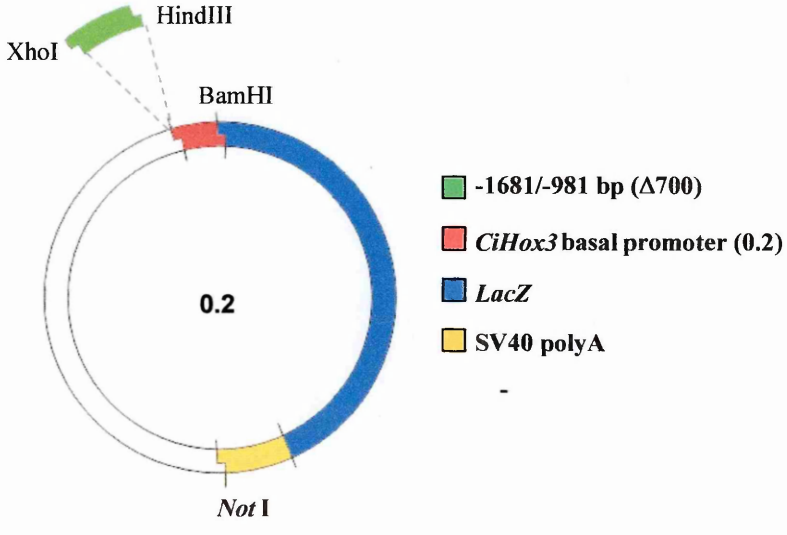


Figure 11

Cloning strategy in 0.2 vector

The insert, indicated in green, is cloned in the *Xho*I-*Hind*III restriction site. In red is indicated the *CiHox3* basal promoter.

PCR screening, in order to have the *Xho*I and *Bam*H1 sites, present in the pCR® II-TOPO® polylinker, in 3'-5' orientation. Then, after digestion, the *Bam*H1 – *Xho*I insert was ligated in the pBSLacZ, digested with the same enzymes, with a 1:3 ratio vector/insert. The promoter fragment -1681/-981 bp was also attached to the heterologous promoter of *CiHox3* by directional cloning of the fragment into the *Xho*I-*Hind*III digested 0.2 construct (**fig. 11**), described from Locascio *et al.*, (1999), to obtain the transgene Δ700/hox3 (**fig. 33**).

The genomic fragments -1681/-1451, -1461/-1221, -1231/-990, -1681/-1221 and -1461/-990 contained in the constructs X/X1, X/X2, X/X3, X1-2 and X2-3 respectively (**fig. 35**), were amplified by PCR from the intermediate construct 700/ TOPO. The oligonucleotides used for amplification were designed according to the sequence of the genomic DNA and contained restriction site *Xho*I at 5' and *Xba*I at 3'. Each PCR amplification reaction was performed as previously described. The amplified fragments, analysed and extracted from 1.5% gel agarose, were ligated in a vector already available in laboratory, containing E1BTATA promoter sequence just upstream *LacZ* reporter gene (**fig. 12**).

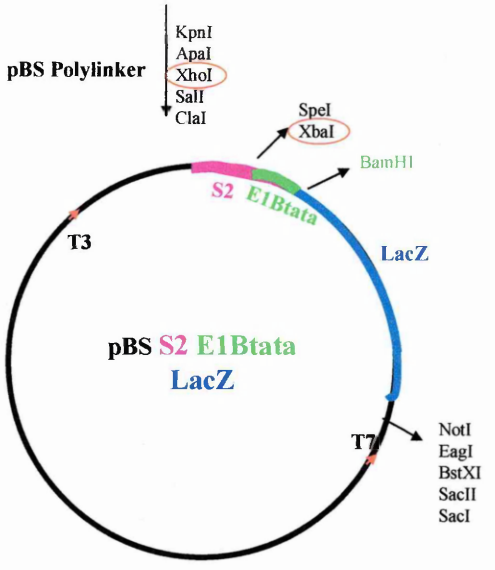


Figure 12

pBS/S2/E1BTATA/LacZ vector

The vector contains the E1BTATA promoter sequence upstream *LacZ* reporter genes

This vector was digested with *Xho*I and *Xba*I enzymes to eliminate S2 promoter fragment that was substituted with the PCR fragments digested with the same restriction

enzymes. The ligase reaction was conducted at 16°C over night with a ratio 1:2 plasmid/insert.

The constructs 1a, 1b, 2a, 2b, 1a1b2a, 1b2a, 1b2a2b containing the genomic fragments -1681/-1571, -1581/-1451, -1461/-1320, -1336/-1221, -1681/-1320, -1581/-1320, -1581/-1221, respectively (**fig. 40**) were prepared in the same way. In the preparation of all the above constructs, the PCR screening methods was used to select the clones containing the inserts. In particular, when PCR was used, in the cases in which the insert was oriented by the presence of two different restriction sites at the ends, two primers internal to the insert were used. In the cases in which, on the contrary, the insert presented the same restriction site at both the ends, two primers, one complementary to the insert and the other complementary to the vector, were used. The plasmidic DNA from the positive clones was prepared and sequenced. The recombinant plasmids were amplified by electroporation as already described in Top10 bacterial competent cells with a transformation efficiency more than $5x10^8$/μg of DNA on average. The bacterial cells were plated and grown at 37°C overnight. Plasmid DNA maxi-preparations were purified on CsCl gradient as previously described.

Preparation of constructs: *5XS2E1b/CiTitf1* and *5XS2E1b/CisFrp1/5*

To prepare the constructs used in the overexpression experiments I used constructs already available in laboratory: pBS/S2E1BTATA/GFP, pBS/*Ci*Bra/*Ci*Titf1. First of all the pBS/5XS2E1BTATA/GFP (**fig. 13**) was digested with *Sal*I to eliminate GFP. The resulted plasmid was run on agarose gel and purified by QiuaQuick purification kit according to manufacture's instruction and ligated.

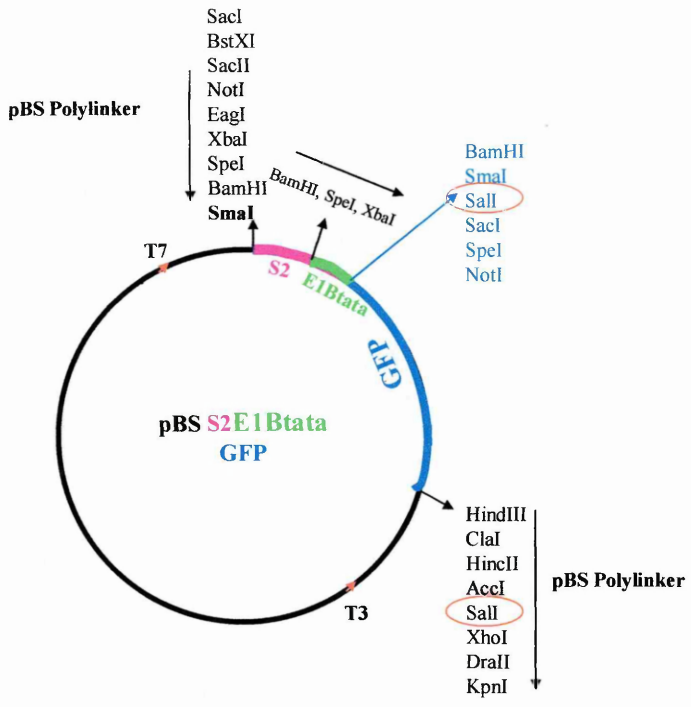


Figure 13

pBS/5XS2E1BTATA/GFP vector map

The vector contains 5XS2 nucleotides upstream E1BTATA

This pBS/5XS2E1BTATA transgene contains the concatameric 5XS2 oligonucleotide upstream E1BTATA promoter. To prepare the construct *5XS2E1b/CiTitf1*, the coding sequence of *CiTitf1* plus SV40 was obtained digesting pBS/*Ci*Bra/*Ci*Titf1 (**fig. 14**) with *Sma*I and *Sac*I. The *Sma*I - *Sac*I insert was then ligated in the pBS/5XS2E1bTATA construct digested with *Sma*I.

Successively, the non cohesive ends were filled-in and ligated. The ligase reaction was conducted at 16°C over night with a ratio 1:2 plasmid/insert.

5XS2/CiTitf1 construct was used as starting vector.

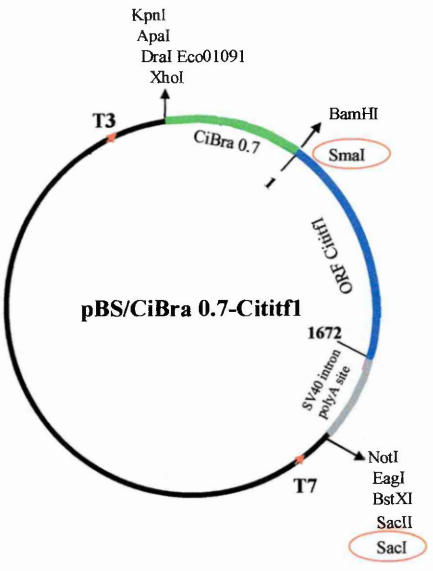


Figure 14

pBS/*Ci*Bra/*Ci*Titf1 vector map (Spagnuolo and Di Lauro, 2002)

The vector contains ORF *Ci*Titf1used to prepare *5XS2E1b/CiTitf1* construct and SV40 polyadenylation signal.

To prepare the construct *5XS2E1b/CisFrp1/5*, the *CisFrp1/5* cDNA clone was digested with *Sma*I and *Xho*I. The *Sma*I - *Xho*I insert was run on agarose gel, purified as above, and ligated to *5XS2E1b/CiTitf1* pre-digested with *Sma*I and *Mlu*I to eliminate *CiTitf1* and to maintain SV40 polyadenylation signal. Successively, the non cohesive ends were filled-in dephosphorilated and ligated. The ligase reaction was conducted at 16°C over night with a ratio 1:2 plasmid/insert.

Preparation of constructs: *5XS2E1bEn/CiTitf1*, *5XS2E1bEn/CiTitf1WF*

To prepare the constructs used in the interference experiments I used constructs pRN3/En-*Ci*Titf1 and pRN3/En-*Ci*Titf1WF already available in laboratory.

To prepare *5XS2E1bEn/CiTitf1* and *5XS2E1bEn/CiTitf1WF* constructs pRN3/En-*Ci*Titf1 (**fig. 15**) and pRN3/En-*Ci*Titf1WF constructs were digested with *BglII* and *EcoRI*.

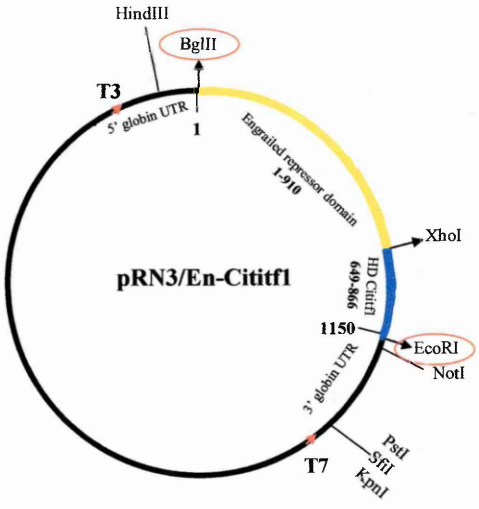


Figure 15

pRN3/En-*Ci*Titf1 vector (Spagnuolo and Di Lauro, 2002)

The vector contains Engrailed repressor domain upstream the Cititif1 homeodomain coding sequence (*CiHD*) used to prepare *5XS2E1bEn/CiTitf1*. Mutagenized versions of pRN3/En-*Ci*Titf1 were produced by changing the WF residues of *Ci*Titf1 into two A: pRN3/En-*Ci*Titf1WF (Spagnuolo and Di Lauro, 2002). By this last vector was obtained En-*Ci*Titf1WF used to prepare *5XS2E1bEn/CiTitf1WF* .

The fragments were run on agarose gel, purified as above and non cohesive ends were filled. Subsequently they were inserted in the *5XS2E1b/CiTitf1* construct digested with *Sma*I and *Mlu*I and filled in. The ligase reaction was conducted at 16°C over night with a ratio 1:2 plasmid/insert. After ligation, the positive clones were selected by PCR screening and corresponding DNA was prepared and sequenced. Plasmid DNA maxi-preparation were purified on CsCl gradient as previously described.

The construct *CiBra/Cititif1*, used in co-electroporation assay, was already available in laboratory (Spagnuolo and Di Lauro, 2002).

Purification on CsCl gradient of DNA constructs

Plasmid DNA was resuspended in 1x TE and for every mg of DNA 1.2 g of CsCl were added. After the addition of 100 µl of EtBr (10 mg/ml) for each DNA/CsCl ml, the samples were transferred in Beckman Quick Seal ultracentrifuge tubes and centrifuged in a VTi-65 rotor 16 hours at 55,000 rpm and at 25°C in a Beckman L8-70M ultracentrifuge. By this technique, circular plasmid DNA was separated by contaminant bacterial DNA and RNA. The separation occurred by virtue of different density acquired by plasmid DNA in the presence of EtBr compared to chromosomal DNA. Two distinct bands were formed on the gradient, the upper one contained nicked bacterial plasmid and chromosomal DNA, the lower one corresponded to supercoiled plasmid DNA. RNA molecules, having higher density, were located at the bottom of the gradient. The band containing the DNA of interest was collected using a 21-gauge needle. The EtBr was removed by adding 1 volume of isoamyl alcohol and by centrifuging at 10,000 rpm 10 minutes. The extraction was repeated several times until EtBr was eliminated from the alcohol phase. Finally, in order to remove CsCl, plasmid DNA was precipitated 15 minutes on ice after the addition of 3 volumes of H_2O and 2 volumes of 100% ethanol and then centrifuged at 10,000 rpm 20 minutes at 4°C. The pellet was washed in 70% ethanol and resuspended in sterile H_2O.

Fertilised eggs electroporation

This technique was used to transfect DNA constructs into fertilised *C. intestinalis* eggs and to observe their expression level in embryos at different stages. Eggs were treated 5-6 minutes with a pH 10 solution containing: 1% thioglycolic acid, 0.05% Pronase E in filtered seawater (FSW). This step permitted to eliminate chorion

and follicular cells surrounding and protecting eggs, which were then fertilised in FSW with spermatozoa collected from 2 or more individuals to avoid self-sterility problems. After 10 minutes, 2-4 rapid washes were carried out to eliminate the exceeding spermatozoa, and then eggs were transferred in a solution containing 0.77 M mannitol and 50-100 μg DNA. Electroporation was executed in Bio-Rad Gene Pulser 0.4 cm cuvettes using Bio-Rad Gene Pulser II with extension capability, at constant 50 V and 500-800 μF, in order to have a time constant between 14 and 20 mseconds.

Embryo fixation and staining

Embryos were allowed to develop at 18-20 °C on 0.9% agarose-coated dishes in FSW until the desired developmental stage, then fixed in 1 % glutaraldehyde in FSW 15 minutes at RT, washed twice with 1x PBS and stained at 30 °C in staining solution (3 mM $K_3Fe(CN)_6$, 3 mM $K_4Fe(CN)_6$, 1 mM $MgCl_2$, 0.1% Tween 20 and 250 μg/ml X-gal in 1x PBS. After incubation, embryos were washed in 1x PBS and transferred on slides for the observation at microscope. The analysis of the different constructs was carried out on the same batch of electroporated embryos, which were stained in parallel and for the same time length. For each construct a minimum of 30 embryos were analysed in at least five different electroporations.

LacZ staining was carried out also by *in situ* hybridisation assay, as described previously, using *LacZ* RNA as riboprobe. For detailed microscopic observations samples were postfixed in 4% paraphormaldehyde and, after dehydration through graded alcohol series, they were incubated in propilene oxide and propylene oxide EPON resin 1:1. Samples were then embedded in EPON resin. After polymerization semithin sections were cut with a Leica Ultracut UCT microtome and were observed under a Zeiss Axiophot microscope.

Results

1 Identification of genes differentially expressed

A cDNA library was prepared by subcloning the subtracted cDNAs corresponding to the RNAs extracted from late neurula stage embryos electroporated with *Ci-Bra/Cititf1* transgene and control embryos electroporated with *Ci-Bra/LacZ*. The library contained almost 50000 cDNA clones; 1000 clones were randomly picked and sequenced on both 5' and 3' regions. The sequences were blasted against various online databases to elucidate the identity of each clone. These databases included the National Center for Biotechnology Information (NCBI), JGI *Ciona* genome database (http://genome.jgi-psf.org/*Ciona*4/*Ciona*4.home.html) and *Ciona intestinalis* cDNA databases (http://ghost.zool.kyoto-u.ac.jp/indexr1.html). Sequence analysis suggested that almost 30% of the clones represented different genes.

Among the sequences indicating genes with a predicted function, I focused my attention on 200 cDNA clones that appeared to encode a broad spectrum of different products, including 40 unknown proteins, 1 cell adhesion protein, 6 cell cycle proteins, 43 enzymes, 14 kinase and kinase-interacting proteins, 13 DNA binding proteins, 2 leucine-rich repeat proteins, 2 cysteine-rich repeats proteins, 19 ATP-GTP-cAMP metal-binding proteins, 42 cellular proteins and 22 RNA-processing proteins.

1.2 Looking for endoderm-specific expression pattern clone

Each one of the 200 cDNAs was used for *in situ* hybridization assays to select those exhibiting expression in endodermal territories.

In order to obtain longer probes, the corresponding clones were picked from the arrayed cDNA *C. intestinalis* collection available in our Laboratory (kindly provided by Prof. Nori Satoh). After DNA preparation, the clones were sequenced to verify the presence of the correct insert. *In situ* hybridization experiments showed that a total of 10 clones (the sequence information is summarized in **Table 2**), were expressed in endodermal territories: three examples are presented in **figure 16**.

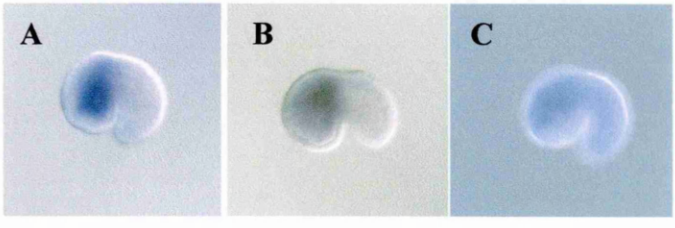


Figure 16

***In situ* hybridization assays results**

Digoxigenin-labeled RNA probes were prepared with cDNA clones isolated via subtractive hybridization. These probes were used for in *in situ* hybridization assays. **A**) Spatial expression, at the early tailbud stage, of a gene encoding an apolipophorins precursor (collection ID GC27g12). **B**), **C**) Spatial expression of two genes showing no sequence similarity with any known factor (collection ID GC27l11 and GC09k04 respectively).

Satoh gene collection ID	Sequence similarity (BLASTP)	Figures
I uknown genes		
GC22p23	Similar to Zinc finger CCHC domain containing protein 4 [*Bos taurus*]	
GC07i16	Similar to CG14331-PA [*Strongylocentrotus purpuratus*]	
GC25h06	putative cellulose synthase [*Ciona intestinalis*]	
GC27l11	hypothetical protein [*Plasmodium falciparum* 3D7]	16 B
GC09k04	Similar to MGC80268 protein [*Xenopus laevis]*	16 C
GC29g24	Similar to RIKEN CDNA [*Homo sapiens*]	
II DNA binding proteins		
GC31f13	Zn-finger protein	
III Cysteine rich repeats proteins		
GC18i18	Cystein-rich protein like	
IV cellular proteins		
GC27g12	Apolipophorin precursor	16 A
GC28p23	Importin-chromosomal region maintenance protein	
GC28d06	Secreted frizzled related protein	22, 23, 24

Table 2

BLASTP analysis of the clones isolated by SSH showing endodermal expression in *C. intestinalis*.

2.1 *CisFrp1/5*- transcript is detected in endodermal cells

One of the genes isolated in this screen coded for a secreted frizzled related protein and was expressed, in the early neurula, in the anterior region of the embryos, which includes the endoderm. As already mentioned, secreted frizzled-related proteins (sFrps) were considered to be absent in invertebrates. The presence of a sFrp in the endodermal territories of a chordate ancestor, together with the still poor information on this gene family, especially in lower organisms, prompted me to carry out a detailed analysis of this gene in *Ciona*, focusing my attention on its expression pattern

throughout embryogenesis, on the structure of its 5'-regulatory region and on its potential relationship to *Cititf1*.

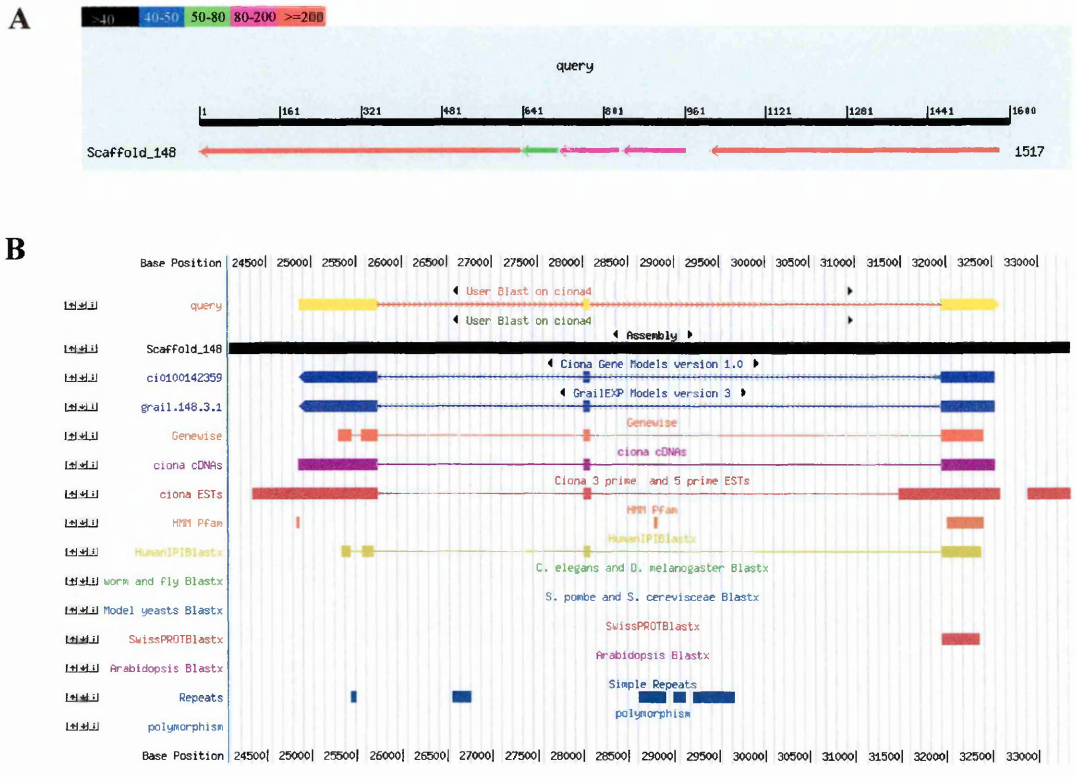


Figure 17

JGI *CisFrp1/5*

A) Result of the BLAST analysis against the JGI genome project. The major homology was found in Scaffold_148 where *Cis*Frp1/5 was annotated. **B**) The Genome Viewer is shown in which the Scaffold track is the black line on the top. Then, the *CisFrp1/5* corresponding JGI-predicted gene is represented in the Gene models track. Finally, nucleotide sequences relevant to the organism of interest are shown. These include ETSs, Pfam Hidden Markov Models search and BLAST data for related organisms.

Figure 17 displays the results of a blust analysis for the clone GC28d06 in the *Ciona* 1.0 genome database. The sequence was found to be located in the Scaffold_148, where it was annotated as *CisFrp1/5* gene.

2.2 *In silico* analysis of *CisFrp1/5* sequence

The complete nucleotide sequence of *CisFrp1/5* cDNA clone revealed an insert length of 1600 bp, containing a 1096 bp open reading frame that starts with an ATG in position 70 and ends with a stop codon in position 1166, followed by 399 bp of untranslated trailer, comprising a non canonical polyadenilation signal (Graber *et al.*, 1999), 29 nucleotides upstream from an $(A)_{21}$ stretch. The reading frame encoded a protein of 365 amino acids, containing the cysteine rich domain (CRD) (shown in light blu in **fig. 18**) at the N-terminal, and the netrin domain (NTR) (shown in yellow in **fig. 18**) at the C-terminal. The putative starting methionine is followed by a signal sequence (underlined in **fig. 18**) with a cleavage site located between residues 34 and 35 (Nielsen *et al.*, 1997). The presence of a signal sequence suggests that *CisFrp1/5* is a secreted protein.

```
   1 ACAAGCTTTAGTGTTACGAGTCTTGGTTTAGTTCTCATGTATAAAAACTATGTTTTTTCA 60
  61 GAGCCCCGAAATGGGATCGTGGATAAAAGGAAACTGTTCTAGTAATGAAGACATTTTTAT 120
                M  G  S  W  I  K  G  N  C  S  S  N  E  D  I  F  I  -

 121 AAGCCTTGCGATATTAATATTATTAGTAAACAATGTTCAGAACTGTTGTGGCAGCCCGAA 180
      S  L  A  I  L  I  L  L  V  N  N  V  Q  N  C  C  G  S  P  K  -

 181 GTTCATGGCGGATGGTACAATGTACGCACGGCCTCCTTCTGTATGTGTCGACATACCGAC 240
      F  M  A  D  G  T  M  Y  A  R  P  P  S  V  C  V  D  I  P  T

 241 GTCACTAAAACTTTGCCGGAACGTAGGTTACACGCGCATGGTGCTACCAAATCTTCTGAA 300
      S  L  K  L  C  R  N  V  G  Y  T  R  M  V  L  P  N  L  L  N

 301 TCATGAGTCTCTCACCGAAGTTGTGCAACAGGCAACAAGCTTTGTTCCTCTGTTGTCAAG 360
      H  E  S  L  T  E  V  V  Q  Q  A  T  S  F  V  P  L  L  S  R

 361 ACACTGCCACGCAAACGTCCAAGTTTTTCTCTGTTCACTTTTTGCACCGGTTTGCATACC 420
      H  C  H  A  N  V  Q  V  F  L  C  S  L  F  A  P  V  C  I  P

 421 AACACCACCCGGTGAATCGATGAATGGCCCGATTCCACCTTGCCGCTCGCTCTGTGAATC 480
      T  P  P  G  E  S  M  N  G  P  I  P  P  C  R  S  L  C  E  S

 481 GGTAGAAAGACGATGCGGCCCTATTATGCTGGAACACGGATTTTCCTGGCCACTCATGCT 540
      V  E  R  R  C  G  P  I  M  L  E  H  G  F  S  W  P  L  M  L

 541 TAACTGCAGCAAGTTCAGCAACGAAGGTCTGTGTGTAAGTCCAGATGTCAATGGAAGCAC 600
      N  C  S  K  F  S  N  E  G  L  C  V  S  P  D  V  N  G  S  T

 601 AACTGCTGAACCACCTTTACCTTCGACTGCACCAAATCCAGTTTGTCCACCTTGCAGAAT 660
      T  A  E  P  P  L  P  S  T  A  P  N  P  V  C  P  P  C  R  I

 661 AGAATTGCAGAGAGACACCTTACTTGATAACTATTGTGCAAGCGAATTTGTTATTAAAAT 720
      E  L  Q  R  D  T  L  L  D  N  Y  C  A  S  E  F  V  I  K  I

 721 CCGGGTCAAGAAAATCAAAAATAAGAAAAGTCGGGGAATTCGAGAAGTTTTTGCCGATAA 780
      R  V  K  K  I  K  N  K  K  S  R  G  I  R  E  V  F  A  D  K

 781 GAAGAAAAGAGTAATTTATAAAAATGGGCCACTCGACAAAAGAGATAATAAAAAAATGAA 840
      K  K  R  V  I  Y  K  N  G  P  L  D  K  R  D  N  K  K  M  K

 841 ACTGTACGTGATCGGTAGAAAGAGATGTACGTGCCCTCAGTTGGACGATGCAGTCACTAA 900
```

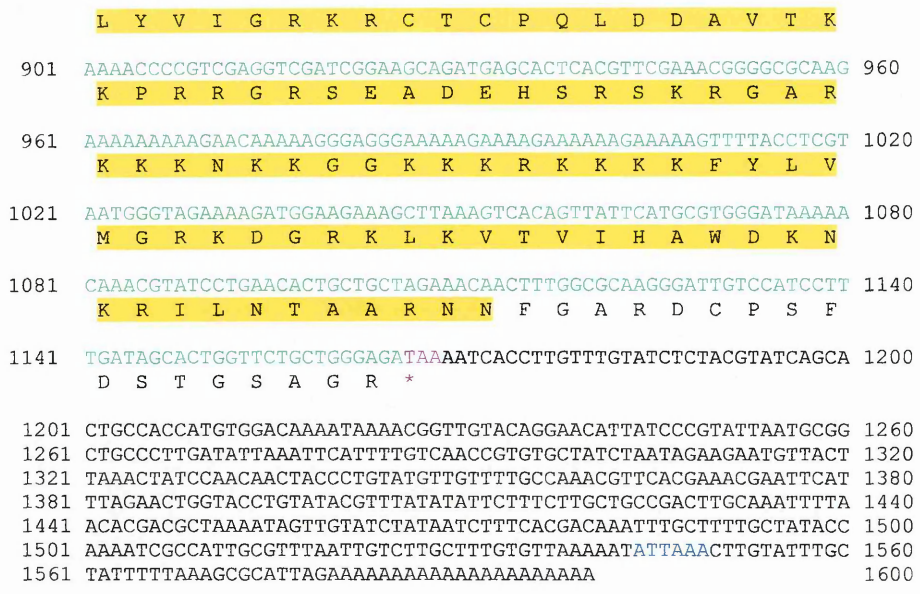


Figure 18

***CisFrp1/5* cDNA sequence and deduced amino acid sequence.**

Coding sequence is represented in green. The nucleotide sequence was confirmed by sequencing both DNA strands. In the sequence the ATG starting codon and the putative starting methionin are bold-face red, the stop codon is violet and the polyadenilation signal, near the 3'end of the mRNA, is shown in blue. A predicted signal sequence is underlined. The CRD domain is shaded in light blue and the NTR domain is shaded in yellow.

CisFrp1/5 aminoacid sequence was analyzed by online Interproscan (European bioinformatics Institute; Zdobnov *et al.*, 2001) and SMART (Simple Modular Architecture Research Tool) programs. These analyses, as shown in **figure 19**, indicated that a Frz cysteine-rich domain spans from amino acids 51-173 (aa light blue in **fig. 19**) while the sequence from amino acids 205-347 constitutes the Netrin-like domain (aa yellow in **fig. 19**) that is also present in tissue inhibitors of metalloproteinases (TIMP).

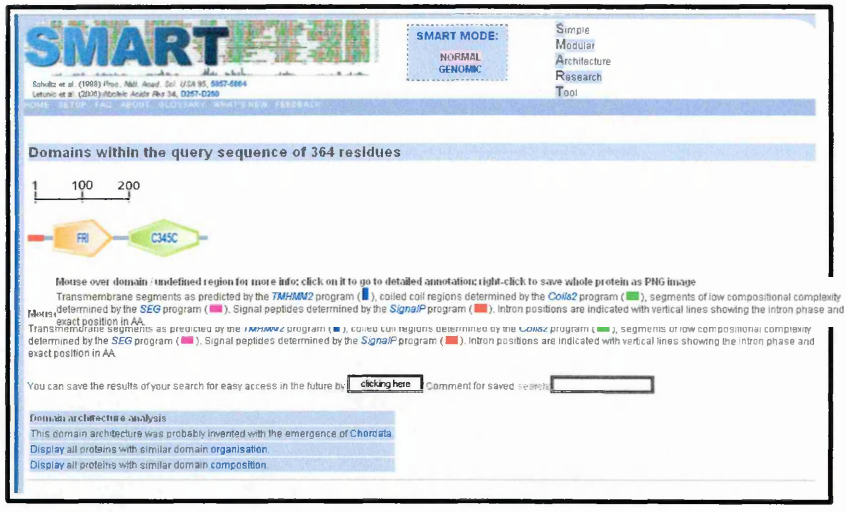


Figure 19

SMART results of *CisFrp1/5* amino acid sequence analysis

Modular architecture analysis showing the domains within the *CisFrp1/5* deduced aminoacid sequence. FRI domain is indicated in orange, NTR domain in green and signal peptide in red.

As expected from data reported in the literature, the 122 amino acids of the FRZ domain (Cystein-rich domain, CRD) are highly conserved and show the highest homology to mammalian sFrp5 and sFrp1 (Hino K. *et al.*, 2003) (**fig. 20**). *CisFrp1/5* does not contain a transmembrane domain or the Lys-Thr-X-X-X-Trp motif, required by frizzled receptors for signalling (Umbhauer *et al.*, 2000), and unusually contains a serine instead of a conserved proline (highlighted in blue in **fig. 20**), 4 residues far from the cysteine 9, at the C-terminal.

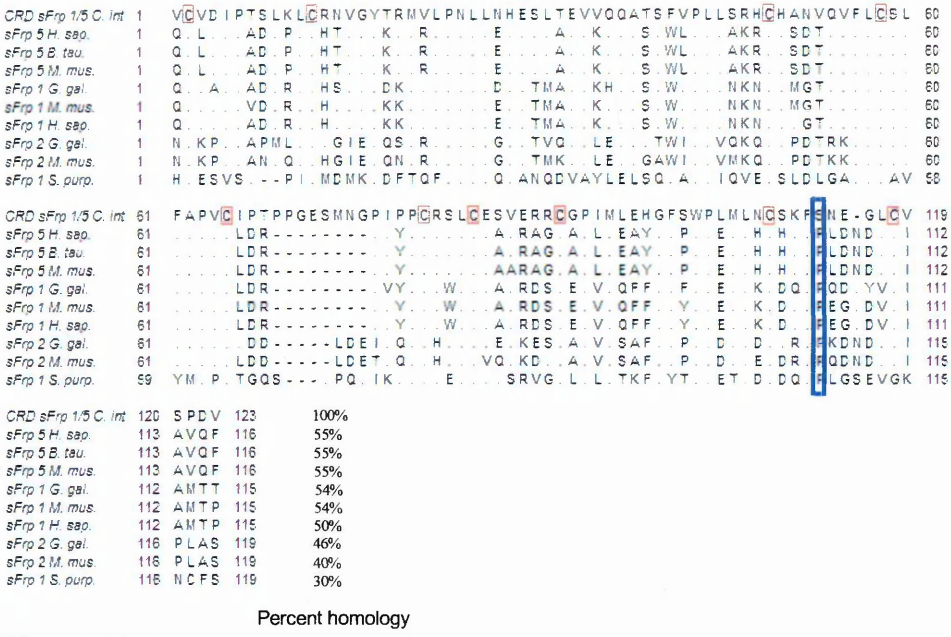


Figure 20 A

Homology of *CisFrp1/5* cystein-rich domain (CRD)

Comparison of *CisFrp1/5* CRD deduced sequence with CRD domain sequences isolated from other species. On the right, percentages of homology are indicated. C. int, *Ciona intestinalis*; H.sap, *Homo sapiens*; B.tau., *Bos taurus*; M.mus., *Mus musculus*; G.gal, *Gallus gallus*; S.purp, *Strongylocentrotus purpuratus*. In red are indicated the ten key conserved cysteines. The blue box, at the C-terminal, highlights the serine residue present in *CisFrp1/5* instead of a conserved proline. Dashes indicate identical amino acids. Dots indicate gaps in the sequence. (GenBank Accession numbers: H.sap sFrp5 014780; H.sap sFrp1 O00546; B.tau sgi|4583384; M.mus sfrp2sgi|6677895; M.mus sfrp5sgi|9055340; G.gal sFrp1sgi|11124735; G.gal sFrp2sgi|6746598; S.purp sFrp1 AF426109).

A molecular phylogenetic tree, shown in **Fig.20 B**, defines four phylogenetically related groups.

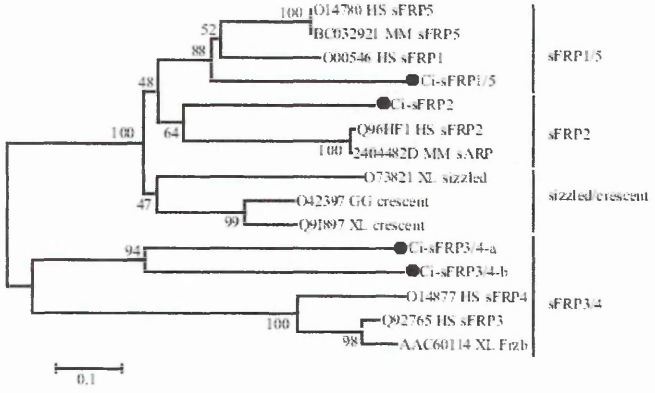


Figure 20 B

Phylogenetic tree of sFRP proteins

The tree is generated by the neighbor-joining method. *Ciona intestinalis* proteins are shown by large black dots. The number beside each branch indicates the percentage of times that a node was supported in 1000 bootstrap pseudoreplications. The sequences used are represented as accession number, abbreviation

of species (HS for human, MM for mouse, GG for Gallus gallus, XL for Xenopus laevis.The scale bar indicates an evolutionary distance of 0.1 amino acid substitutions per position

CisFRP1/5 was assigned to the first group including mammalian sFRP1 and sFRP5. Therefore, at least three ancestral genes were encoded in the genome of the last common ancestor of ascidians and vertebrates; one for sFRP1/5, one for sFRP2, and one for sFRP3/4.

The BLAST analysis against Ciona 1.0 database (Dehal *et al.*, 2002) allowed to predict the *CisFrp1/5* gene structure (**fig. 21**). The gene, spanning 7660 nt, is split in three exons of 571 bp, 75 bp and 869 bp respectively, separated by two introns, one 3877 bp and the other 2268 bp long. This indicates that *CisFrp1/5* has a genomic organization similar to that of mammalian *sFrp1* and *sFrp5*, with the three exons comparable in length (Jones *et al.*, 2002). Interestingly, the structural organization of *CisFrp1/5*, *sFrp1* and *sFrp5* is different from that predicted for the other family members, *sFrp3* and *sFrp4*, which are composed by six exons (Jones *et al.*, 2002). Hence the structural similarity together with their sequence homology is consistent with a relationship between the three genes.

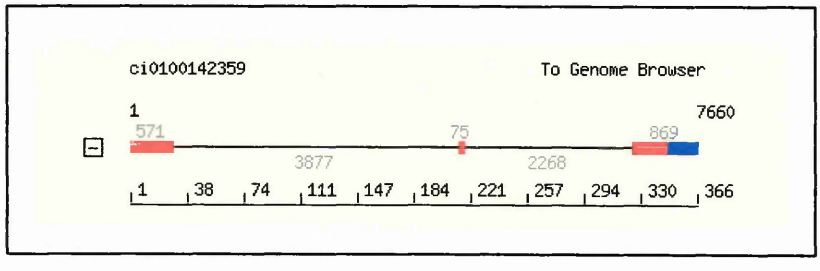


Figure 21

***CisFrp1/5* gene structure**

In red are indicated the three exons of 571 bp, 75 bp and 869 bp. In blue is indicated the 3'UTR.

2.3 Whole mount *in situ* hybridization of *CisFrp1/5*

In situ hybridization experiments indicated that *CisFrp1/5* was expressed at the beginning of gastrulation in the blastomeres precursors of the central nervous system

and of the mesenchyme (**fig. 22 A** and **B**). At the early neurula stage (**fig. 22 C**) the signal, that become more evident at a later stage (**fig. 22 D** and **E**), is confined in the anterior part of the embryos, that includes endoderm and the surrounding ectodermal territories.The expression in the endoderm has been confirmed by semi-thin transverse section at the level of the anterior part of the embryos (**fig. 22 F**).

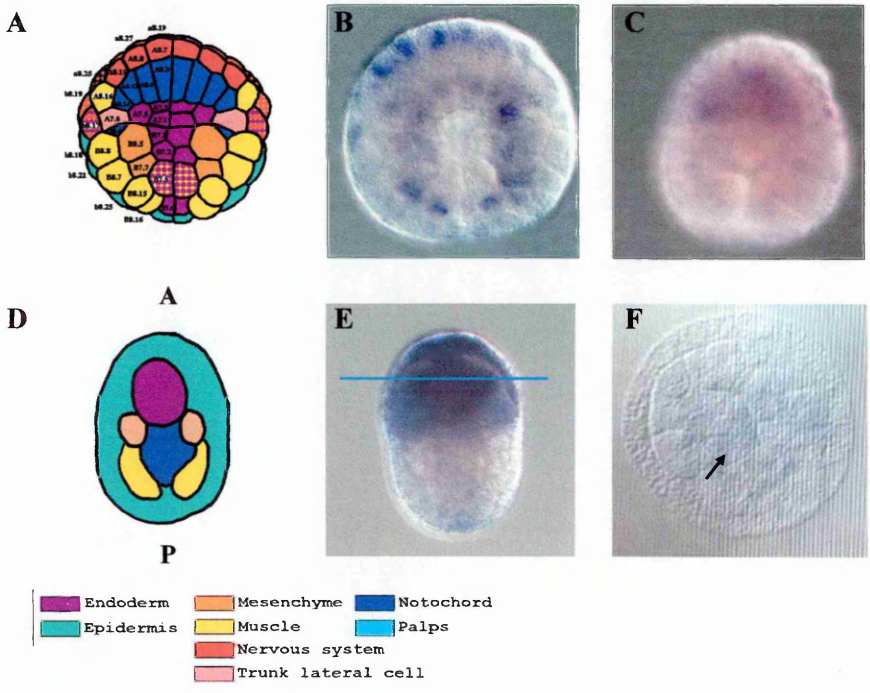


Figure 22

Spatial expression of *CisFrp1/5* at 110 cell, early neurula and neurula stages.

A) Schematic drawing and **B**) vegetal view of a 110 cell-stage embryo. Hybridization signals are visible in the blastomeres that are shown schematically in A. **C**) Ventral view of an embryo at early neurula stage showing expression in the anterior region. **E**) An embryo at neurula stage from dorsal view: intense expression of *CisFrp1/5* is present in the endoderm (indicated in violet in the scheme in D) and in the surrounding ectodermal territories (indicated in green in the scheme in D). **F**) Semi-thin transverse section

at the level indicated in E showing expression in the endoderm cells indicated by a black arrow. The colour code for the drawings A and D are indicated on the bottom left. A: anterior, P: posterior.

Specific staining in the anterior region persisted at the tailbud stage (**fig. 23 A** and **B**). Endoderm expression is corfirmed by semi-thin longitudinal section through the anterior posterior axis of the hybridized embryos (**fig. 23 C**).

Figure 23

Spatial expression of *CisFrp1/5* at tailbud stage.

B) Lateral view of a tailbud stage embryo. Hybridization signal is present in the anterior part of the embryo including endoderm, shown schematically in yellow in **A**. **C**) Semi-thin longitudinal section through the anterior posterior axis of hibridized tailbud stage embryo. Epidermis (blue), muscle (pink), mesenchyme (red), notochord (orange), central nervous system and nerve cord (green). A: anterior, P: posterior, D: dorsal, V: ventral.

At middle (**fig. 24 B**) and late larval stages (**fig. 24 C** and **D**) the transcript appeared in the brain vescicle and in a subset of endodermal cells surrounding ventrally the brain vescicle. No signal was detected with sense probe (data not shown).

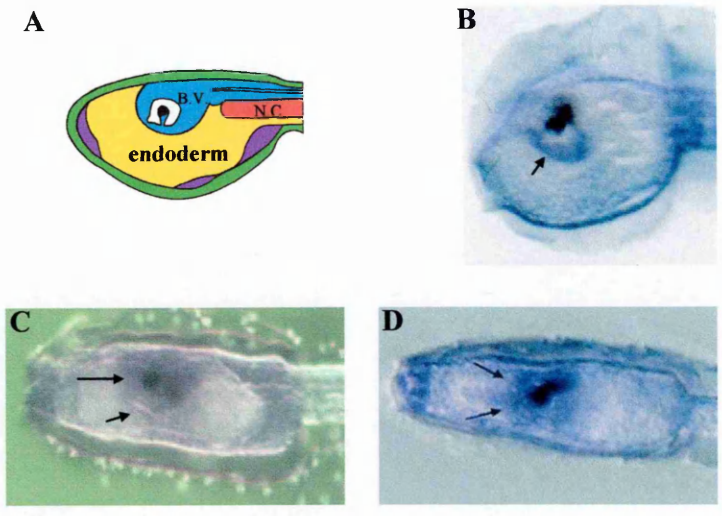


Figure 24

Spatial expression of *CisFrp1/5* at larval stage.

A) A diagrammatic drawing of the tissue organization of a larval stage embryo. Dorsal is on the top. **B)** Lateral view of an embryo at middle tailbud and **C)-D)** swimming larval stage. The arrows show hybridization signals in a subset of endodermal cells surrounding ventrally the brain vescicle. Endoderm (yellow), epidermis (green), mesenchyme pockets (violet), brain vescicle (light blue), notochord (red).

3 *CisFrp 1/5* gene promoter

3.1 *In vivo CisFrp 1/5* gene promoter analysis

In order to study the regulatory mechanisms underling the expression of *CisFrp1/5*, and to identify the *cis*-acting elements required for *CisFrp1/5* expression, I focused my attention on the promoter region. A 5' genomic sequence was amplified by PCR reaction on *C. intestinalis* genomic DNA, using the most suitable oligonucleotides designed on the basis of the genomic sequence available on JGI. Initially I assayed a fragment of 1616 bases, extending from position -1681 to position -65 from the putative ATG starting codon, inserted upstream from *LacZ* reporter gene (**fig. 25**). This construct, named 4/7, was introduced into *C. intestinalis* fertilized eggs via electroporation.

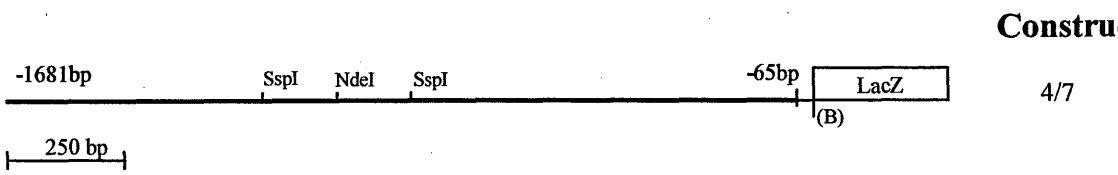


Figure 25

Diagram of 4/7 construct

Diagrammatical representation of the construct 4/7 containing the PCR amplified genomic fragment (-1681 bp to -65 bp) cloned upstream from the *LacZ* reporter gene. The restriction sites (*Ssp*I and *Nde*I) are also indicated.

The embryos were allowed to develop until neurula and tailbud stages, then fixed and assayed either by *LacZ* RNA *in situ* hybridization, or for β-galactosidase activity by X-gal staining. Since no signal was identified, even after long staining times (more than one week), I deeply analyzed the amplified sequence and found that this genomic region, probably, lacked a canonical TATA box that could, instead, be present, at the position -31, in the 65 bases upstream from the ATG that were not included in the fragment 4/7. Therefore I decided to add, at the 3' end of the promoter region, a minimal TATA box already tested in our laboratory (data not published). This consists of the Epstein Barr virus (E1bTATA) minimal promoter (Leong *et al.*, 1998; Parks *et al.*, 1988; Poleev *et al.*, 1995) that is transcriptionally inactive *per se*, but works finely in *Ciona*. E1bTATA sequence is able to amplify the signal driven by the elements cloned upstream from it.

The E1bTATA box was inserted at the 5' end of the oligonucleotide 7 that starts in position -65, as shown in **figure 26**. Oligonucleotide 7E1b was then used, together with oligonucleotide 4 (that starts from position -1681), in a PCR amplification reaction, to add E1bTATA box at 3' end of the 4/7 genomic fragment (see Methods).

At neurula (**fig. 28 A** and **C**) and taibud stages (**fig. 28 E** and **G**), *LacZ* mRNA was present in the anterior region of the embryo, that includes external ectodermal and internal endodermal territories. Embryos were then sectioned in order to highlight the internal endodermal expression. The analysis of semi-thin transverse sections, at level of the anterior part of the embryo, at neurula (**fig. 28 B**) and tailbud stages (**fig. 28 F**) and of semi-thin longitudinal sections, through the anterior posterior axis of the embryo, at neurula (**fig. 28 D**) and tailbud stages (**fig. 28 H**) revealed the presence of a very strong signal in the most anterior endodermal cells. Moreover, almost 10% of the embryos showed also *LacZ* RNA staining ectopically in cells of the trunk mesenchyme.

I submitted the sequence of this genomic fragment to the TRANSFAC database and MatInspector V.6.0 (http://www.genomatix.de/shop/index.html) program, to look for possible binding sites recognised by CiTTF-1. From *in silico* analisys, I identified two CiTTF-1 binding sites at -522 bp (tgtct**CAAG**ttggtt) and -265 bp (cttttt**CAAG**ttttta) upstream from the ATG starting codon; they are situated on the 5'-3' strand. To verify their putative involvement in *CisFrp1/5* regulation, I prepared three constructs in which either both or the single CiTTF-1 two putative binding sites were mutated (**fig. 29a**).

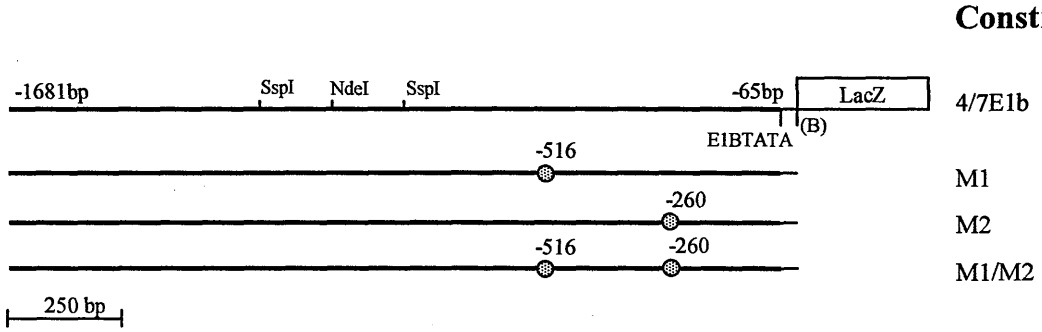


Figure 29a

Diagram of M1, M2 and M1/M2 constructs

The wild type promoter sequence is contained in 4/7E1b construct. The mutations are indicated as green circles and were obtained by site-directed mutagenesis, as detailed in Methods. Mutants sequences were inserted upstream from the *LacZ* reporter gene and electroporated in fertilized eggs. On the right side, the names of the constructs are indicated.

In particular in M1 construct three nucleotides were mutated in the *Cititf1* core sequence at -516 bp (CAA >AGG) while in M2 construct three different nucleotides were mutated in the *Cititf1* core sequence at -260 bp (AAG >CCT). The M1/M2 contruct contained both mutations. All these constructs were analyzed by electroporation into *C. intestinalis* embryos and by *LacZ* mRNA *in situ* hybridization that gives a stronger signal, especially at the neurula stage, compared with the X-Gal staining, probably due to the time required for protein synthesis and LacZ protein accumulation. The construct M1 (**fig. 29b A**), M2 (**fig.29b B**) and M1/M2 (**fig. 29b C**) drove reporter gene expression in the same territories found for the 4/7E1b transgene, indicating that, most probably, these binding sites are not directly involved in *CisFrp1/5* specific promoter activity.

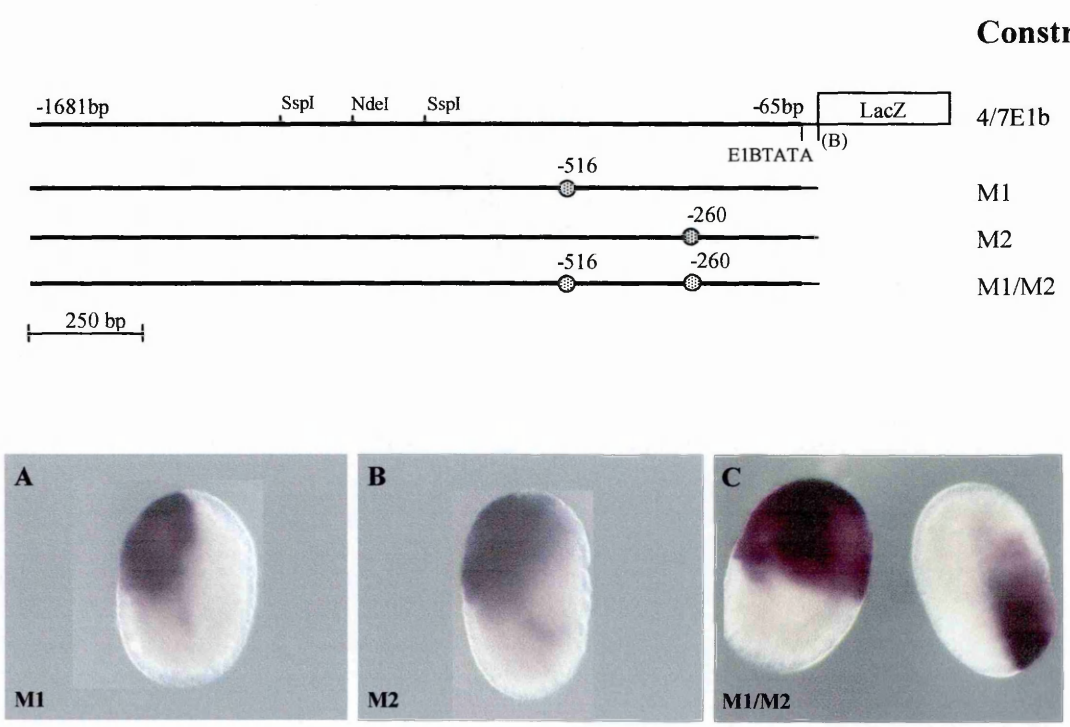


Figure 29b

Expression of the M1, M2 and M1/M2 constructs in electroporated *Ciona* embryos by *LacZ RNA in situ* hybridization

On the top is shown a scheme of the reporter constructs used for electroporation. On the bottom (**A-C**) dorsal view of neurula stage embryos electroprated respectively with M1, M2 and M1/M2 constructs. *LacZRNA* staining is visible in the same territories obtained with the 4/7E1b construct (**fig. 28 A** and **C**).

Consequently, in order to narrow the enhancer element(s) contained in the construct 4/7E1b, I prepared two other constructs, truncated at the 3' or 5' ends, using the restriction site *Nde*I indicated in **fig. 30.**

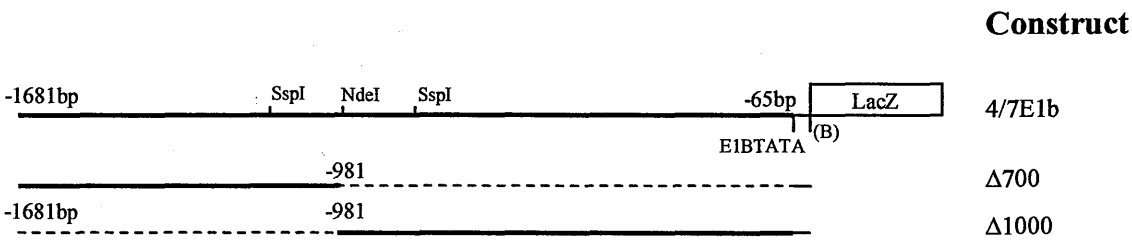


Figure 30

Summary of *CisFrp1/5* transgene constructs

Diagram of different 5' *CisFrp1/5* genomic regions that were inserted upstream from the *LacZ* reporter gene to obtain plasmids for electroporation into *C. intestinalis*. The promoter fragments were obtained by digestion with the indicated restriction enzyme *Nde*I. On the right side, the names of the constructs are indicated.

Introduction of Δ700 construct resulted in the development of embryos where the expression of *LacZ* mRNA reproduced the expression pattern of the whole genomic fragment 4/7E1b at neurula (**fig. 31 A**) and tailbud stages (**fig. 31 C**). These data were confirmed by transverse sections of the anterior part of embryos at neurula (**fig. 31 B**), and tailbud (**fig. 31 D**) stages.

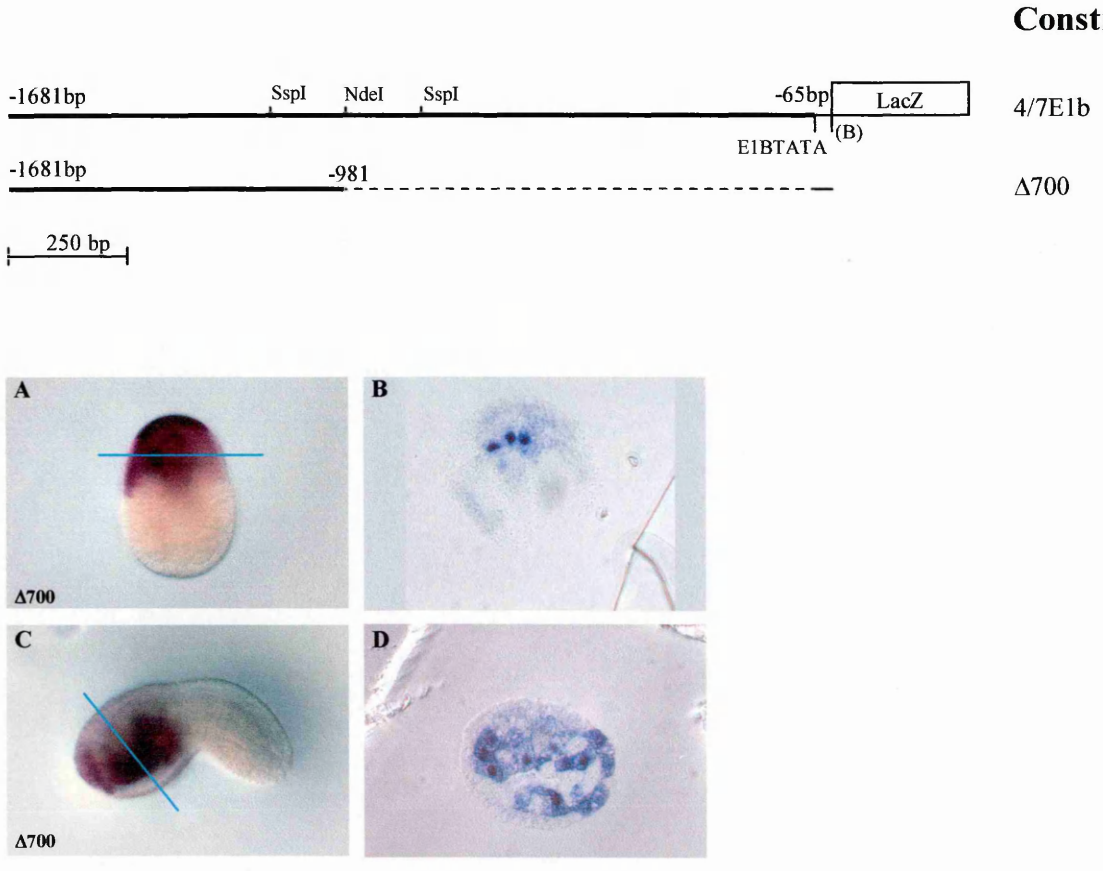


Figure 31

Expression of the Δ700 construct in electroporated *Ciona* embryos by *LacZ RNA in situ* hybridization

Schematic representation of the constructs on the top; electroporated embryos on the bottom. **A)** Dorsal view of a neurula stage embryo and corresponding semi-thin trasverse section **B)** at the level of the anterior part, indicated by a blu line in **A**. **C)** Lateral view of a tailbud stage embryo and corresponding semi-thin transverse section **E)** at the level of the anterior part of the embryo, indicated by a blu line in **C**. *LacZRNA* staining is visible in the same territories obtained with the 4/7E1b construct (**fig. 28 A** and **G**).

In contrast, embryos electroporated with the construct Δ1000 showed only an ectopic *LacZ* RNA staining in mesenchyme cells at the neurula and tailbud stages (**fig. 32 A** and **B**).

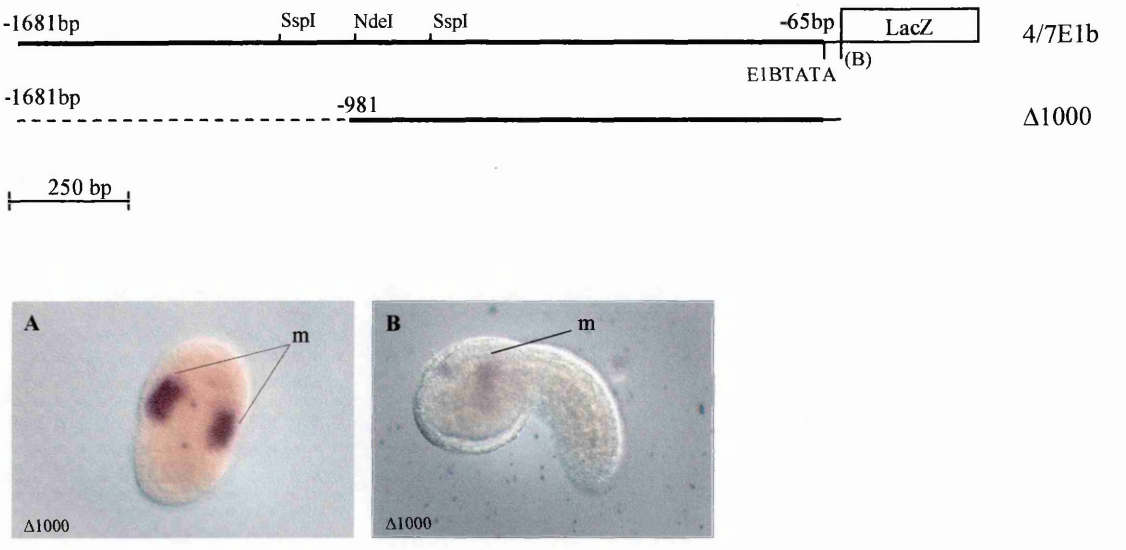


Figure 32

Expression of the Δ1000 construct in electroporated *Ciona* embryos by *LacZ RNA in situ* hybridization

Schematic representation of the constructs on the top; electroporated embryos on the bottom. **A)** Dorsal view of a neurula stage embryo. **B)** Lateral view of a tailbud stage embryo. *LacZ RNA* staining is present, ectopically, in mesenchymal cells.

The results I obtained so far permitted the identification of a promoter region, extending from -1681 to -981, that appeared to reproduce the expression of the endogenous gene: namely the anterior part of the embryo, including the endoderm, and the surrounding ectodermal territories. Since the intensity of the *LacZ* signal and the onset of the expression were comparable to that obtained using the entire promoter fragment (4/7E1b construct), I may suggest that this region contained the *cis*-regulatory information required for the expression of the *CisFrp1/5* gene.

To further prove that the regulatory region I identified was sufficient for this anterior specific expression, I inserted the fragment, extending from -1681 to -981, upstream from the 0.2kb fragment, corresponding to the basal promoter element of

CiHox3 (construct **fig. 33**), that by itself is transcriptionally inactive because it includes only the putative TATA and CAAT boxes for basal transcription machinery of *CiHox3* (Locascio *et al.*,1999). The composite *CisFrp1/5*Δ700/*CiHox3*/*LacZ* construct was found to direct reporter expression with a pattern identical to that obtained with construct 4/7E1b at the neurula (**fig. 33 A**) and tailbud stages (**fig. 33 B**).

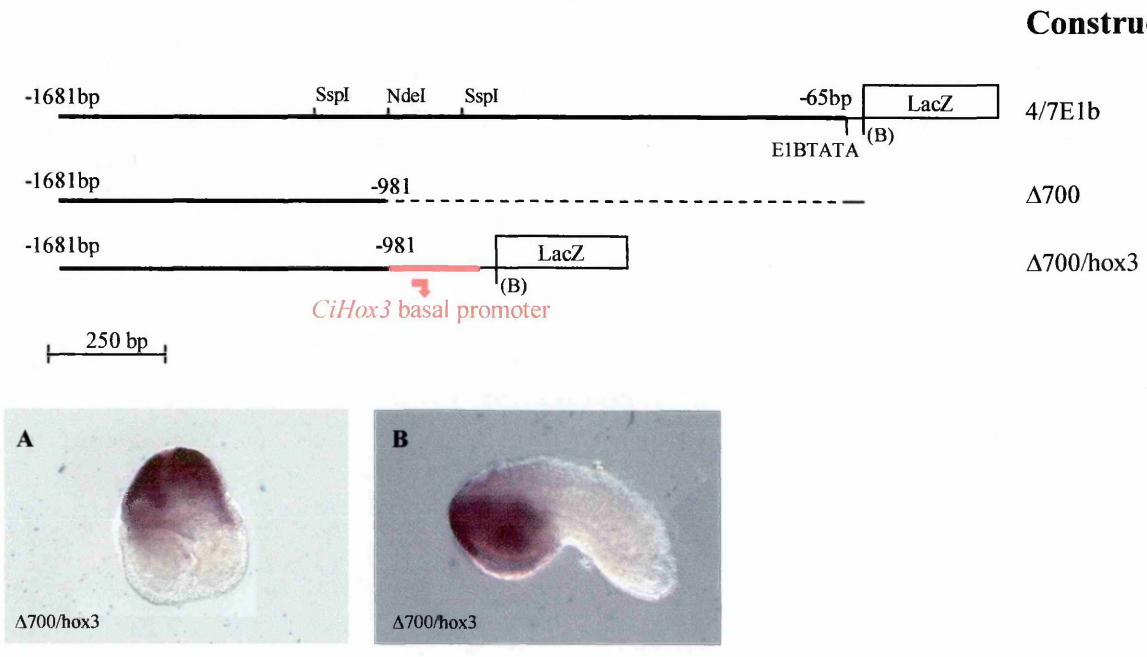


Figure 33

Expression of the Δ700/hox3 construct in electroporated *Ciona* embryos by *LacZ RNA in situ* hybridization

Schematic representation of the constructs on the top; electroporated embryos on the bottom. **A)** Ventral and **B)** lateral view of a tailbud stage embryo. *LacZ RNA* staining is visible in the anterior part of the embryo with a pattern identical to that obtained with construct 4/7E1b at the neurula (**fig. 28 A**) and tailbud stages (**fig. 28 C**).

Based on these results, summarized in **figure 34**, I focused my attention on the sequence extending from -1681bp to -981bp.

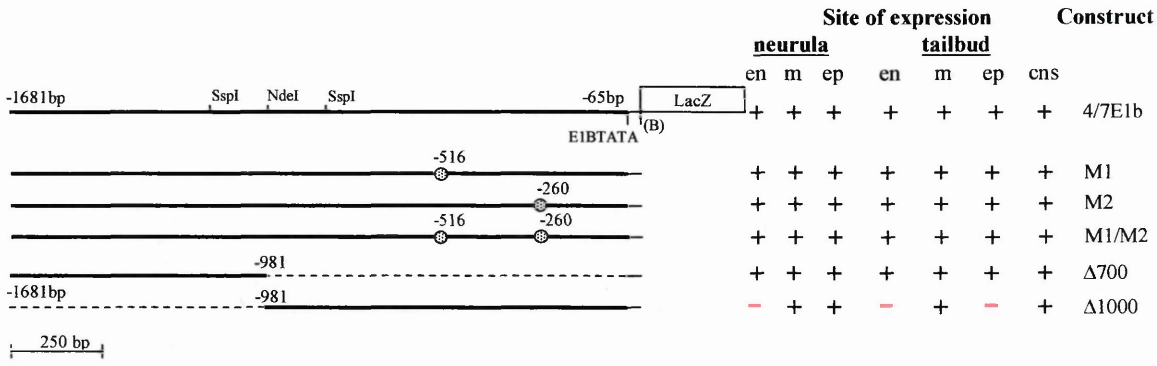


Figure 34

Summary of the results obtained with *CisFrp1/5* transgene constructs

On the left the constructs are schematized. For all the constructs the stage and the territories of reporter gene expression in the embryos are reported on the right.

For this purpose, I prepared, by PCR amplification, five constructs. Three of them were obtained by dividing the region extending from -1681bp to -981bp in three overlapping fragments (X/X1, X/X2, X/X3), long almost 230 bases each (**fig. 35**). Two more constructs were made by joining the fragments two by two (X1-2, X2-3).

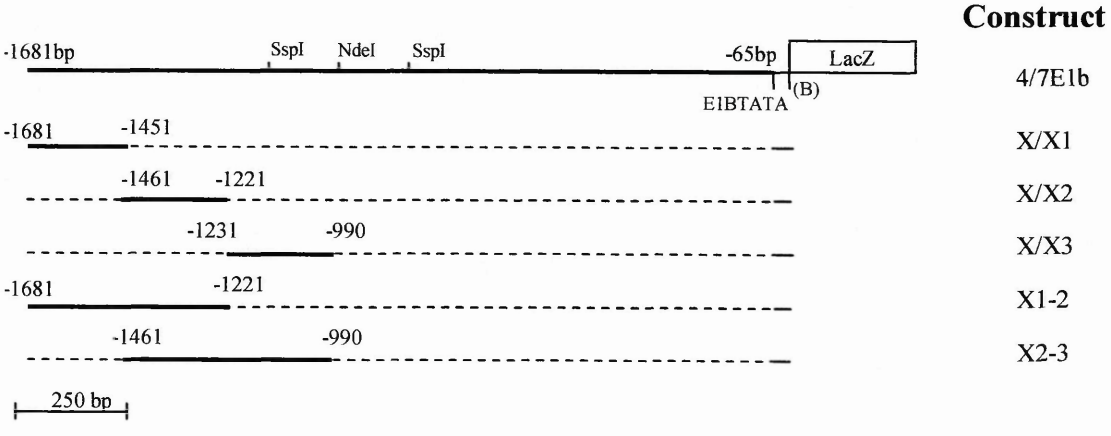


Figure 35

Diagram of the different *CisFrp1/5* promoter regions that were inserted upstream from the *LacZ* reporter gene to obtain plasmids for electroporation into *C. intestinalis* eggs. Constructs names are indicated on the right. The promoter fragments were obtained by PCR, as detailed in Methods.

The introduction of construct X/X1, which contains the promoter sequence from -1681bp to-1451bp, resulted in the development of embryos showing *LacZ* mRNA expression in the anterior region (**fig. 36**) with a staining intensity weaker compared to the construct 4/7E1b.

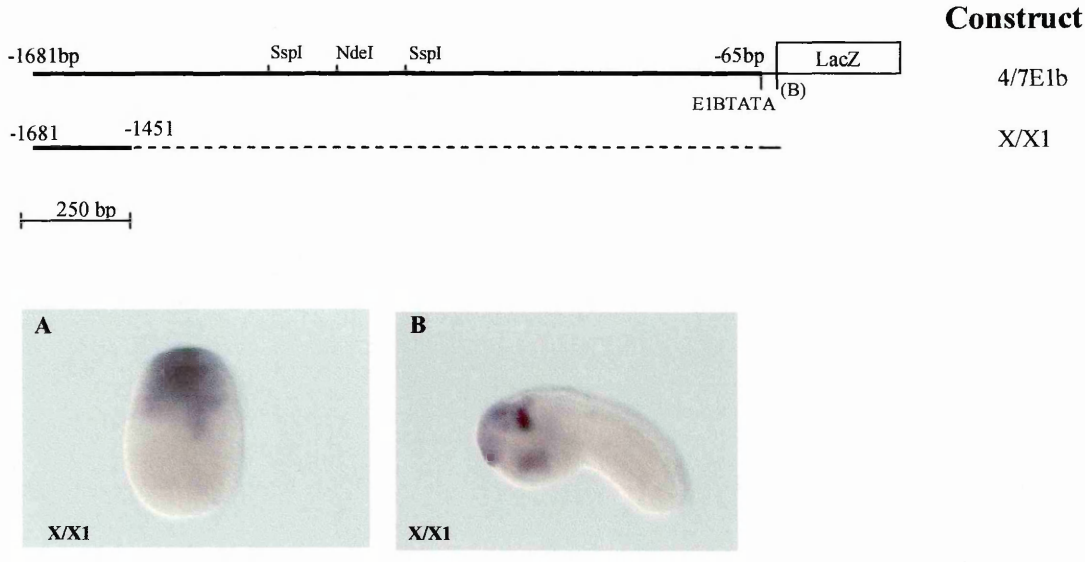


Figure 36

Expression of the X/X1 construct in electroporated *Ciona* embryos by *LacZ RNA in situ* hybridization

Schematic representation of the constructs on the top; electroporated embryos on the bottom. **A)** Dorsal view of a neurula stage embryo. **B)** Lateral view of a tailbud stage embryo *LacZ RNA* staining is visible in the anterior part of the embryo including endoderm, epidermis, with a pattern identical, but weaker, compared to that obtained with construct 4/7E1b (**fig. 28 A**).

In addition the number of stained embryos in the anterior part resulted decreased to 60% compared to the control plasmid 4/7E1b.

Electroporation of construct X/X2, which contains the sequence extending from -1461bp to -1221bp, produced a minority of transgenic embryos, about 30%, showing *LacZ* mRNA staining in the anterior part of the embryo. The remaining embryos presented *LacZ* mRNA expression in the mesenchyme cells (**fig. 37 A** and **B**).

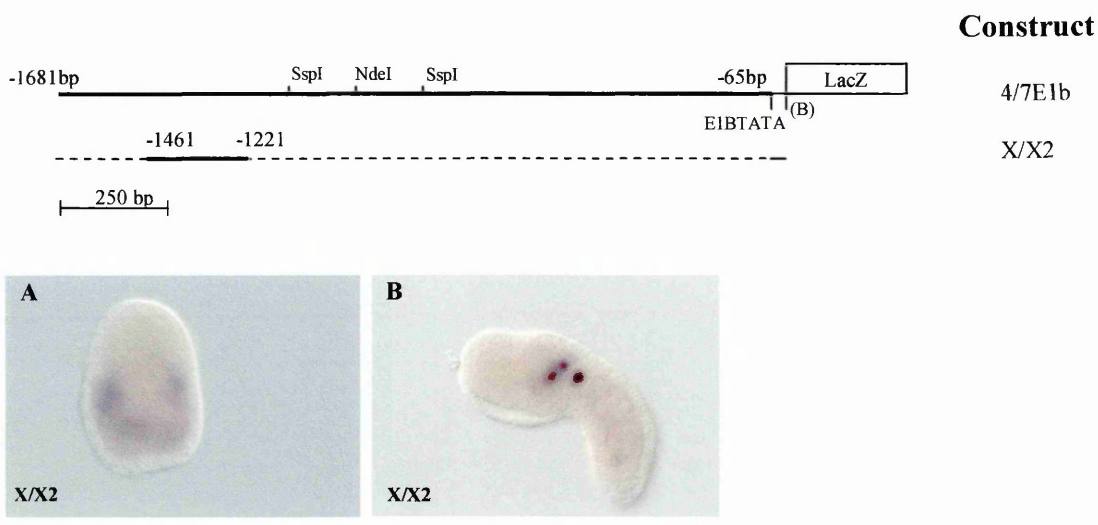


Figure 37

Expression of the X/X2 construct in electroporated *Ciona* embryos by *LacZ RNA in situ* hybridization

Schematic representation of the constructs on the top; electroporated embryos on the bottom. **A**) Ventral view of a neurula stage embryo. **B**) Lateral view of a tailbud stage embryo. *LacZ RNA* staining is present in both stage, ectopically in mesenchymal cells.

Electroporation of transgene X/X3, which contains the sequence from -1231bp to -990 bp, resulted in the development of embryos with no *LacZ* mRNA expression or with a faint signal ectopically in mesenchyme cells both at the neurula (**fig. 38 A**) and tailbud stages (**fig. 38 B**).

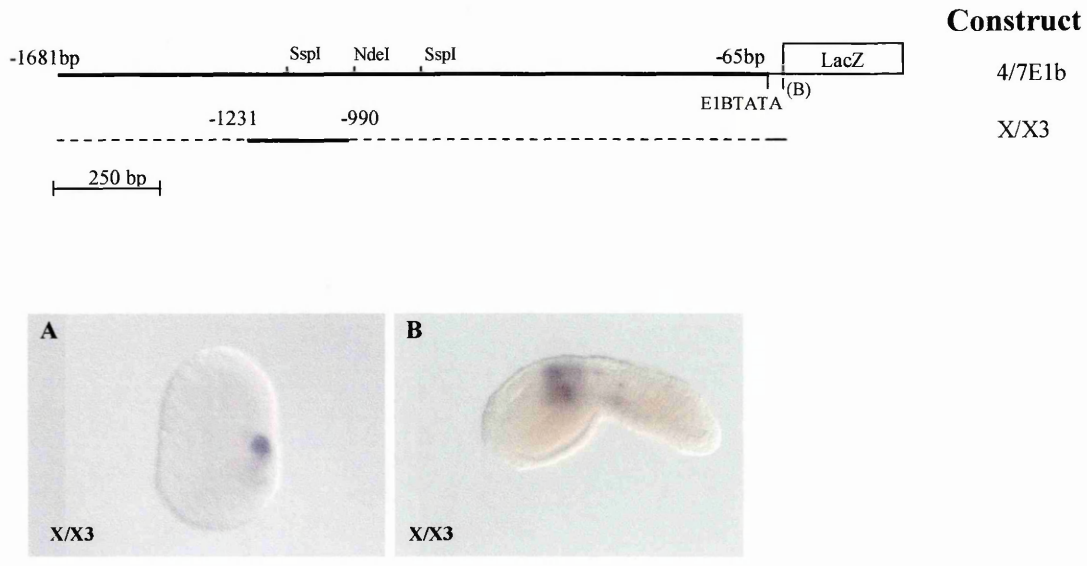


Figure 38

Expression of the X/X3 construct in electroporated *Ciona* embryos by *LacZ RNA in situ* hybridization

Schematic representation of the constructs on the top; electroporated embryos on the bottom. **A**) Dorsal view of a neurula stage embryo. **B**) Lateral view of a tailbud stage embryo. *LacZ RNA* staining is present in both stage, ectopically in mesoderm cells.

These results suggested that the region from -1681bp to -1221bp (that included X/X1 and X/X2) contained the elements that could cooperate to an efficient expression of *CisFrp1/5* gene in the anterior region.

In fact the construct X1-2, extending from -1681bp to -1221bp, reproduced the same pattern observed for construct X/X1, at the neurula (**fig. 39 A**) and tailbud stages (**fig. 39 B**). Moreover the number of the stained embryos in the anterior region resulted increased to 80% compared to X/X1 stained embryos.

In contrast the construct X2-3 was unable to drive any expression in anterior region and only 10% of the embryos presented an ectopic *LacZ* mRNA signal in mesenchyme cells at the neurula (**fig. 39 C**) and tailbud stages (**fig. 39 D**).

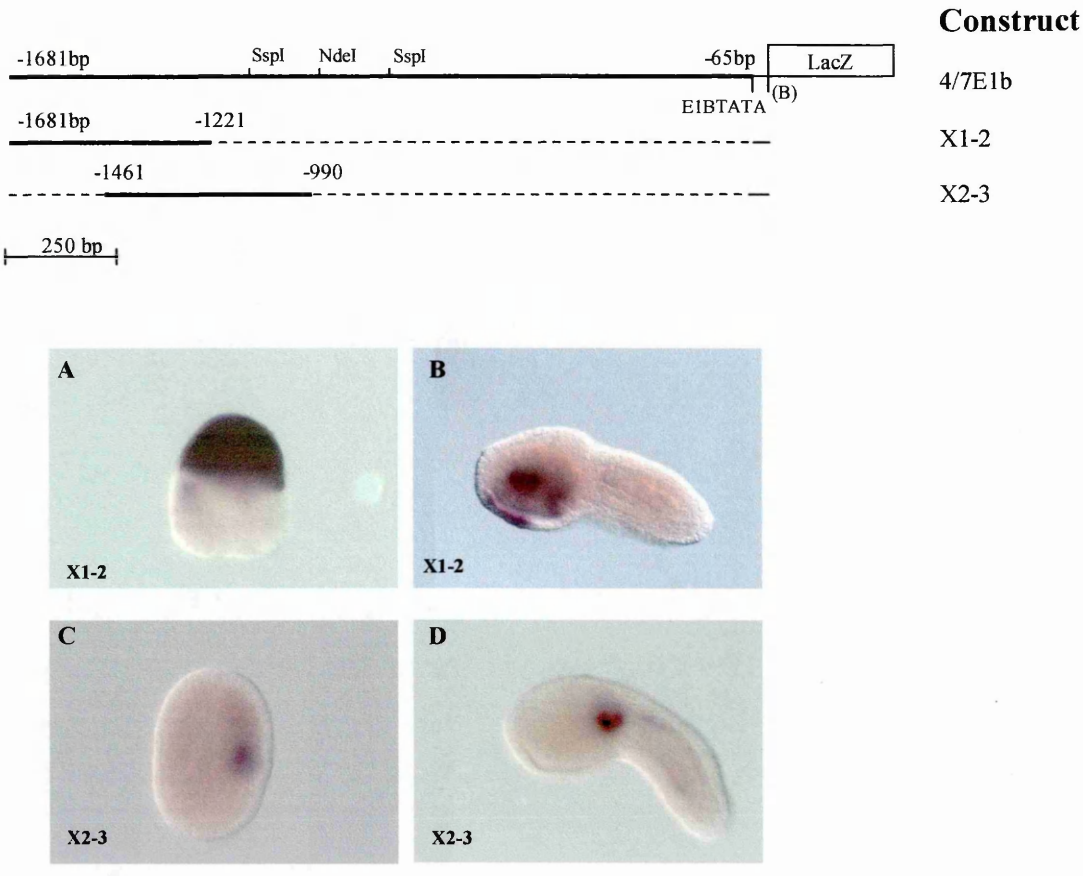


Figure 39

Expression of the X1-2 and X2-3 constructs in electroporated *Ciona* embryos by *LacZ RNA in situ* hybridization

Schematic representation of the constructs on the top; electroporated embryos on the bottom. **A**) Dorsal view of a neurula stage embryo and **B**) lateral view of a tailbud stage embryo electroporated with X1-2 construct. *LacZ RNA* staining is visible in the anterior part of the embryos including endoderm **C**) Dorsal view of a neurula stage embryo and **D**) lateral view of a tailbud stage embryo electroporated with X2-3 construct.. *LacZ RNA* staining is present in both stage, ectopically in mesenchymal cells.

The subsequent analysis was, therefore, focused on the detailed characterization of the fragment extending from -1681bp to -1221bp. For this purpose, this region was divided in four fragments (1a, 1b, 2a, 2b) of about 100 bases each, that were amplified by PCR and tested separately or in different combinations (1a1b2a, 1b2a, 1b2a2b) (**fig. 40**).

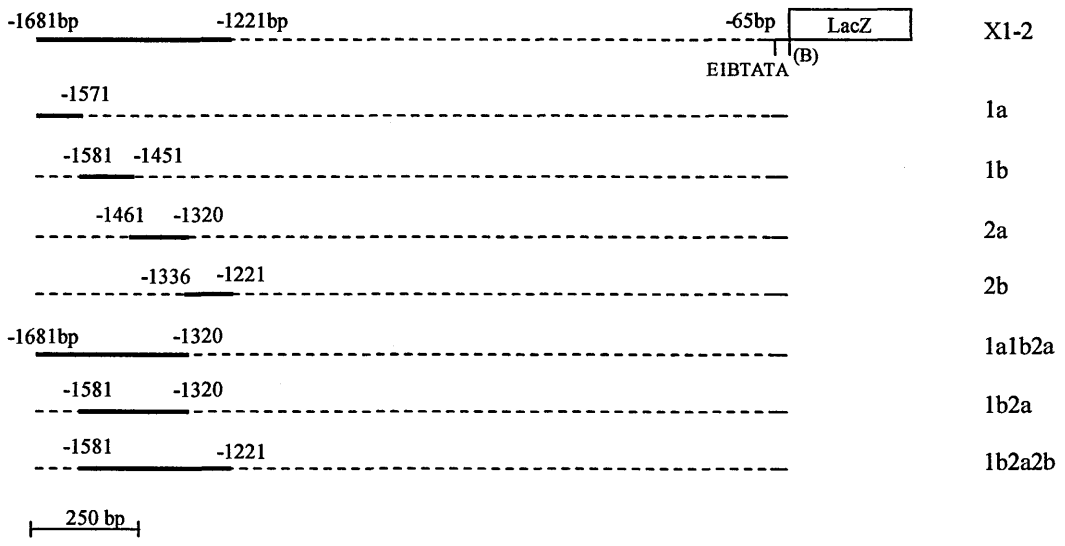


Figure 40

Summary of *CisFrp1/5* constructs

Diagram of the different *CisFrp1/5* promoter sequences that were inserted upstream from the *LacZ* reporter gene to obtain plasmids for electroporation into *C. intestinalis* eggs. Constructs names are indicated on the right. The promoter fragments were obtained by PCR, as detailed in Methods.

Introduction of constructs 1a, 2a, 2b, resulted in the development of embryos with an ectopic expression in mesenchyme cells both at the neurula (**fig. 41 A, C** and **E**) and tailbud stages (**fig. 41 B, D** and **F**).

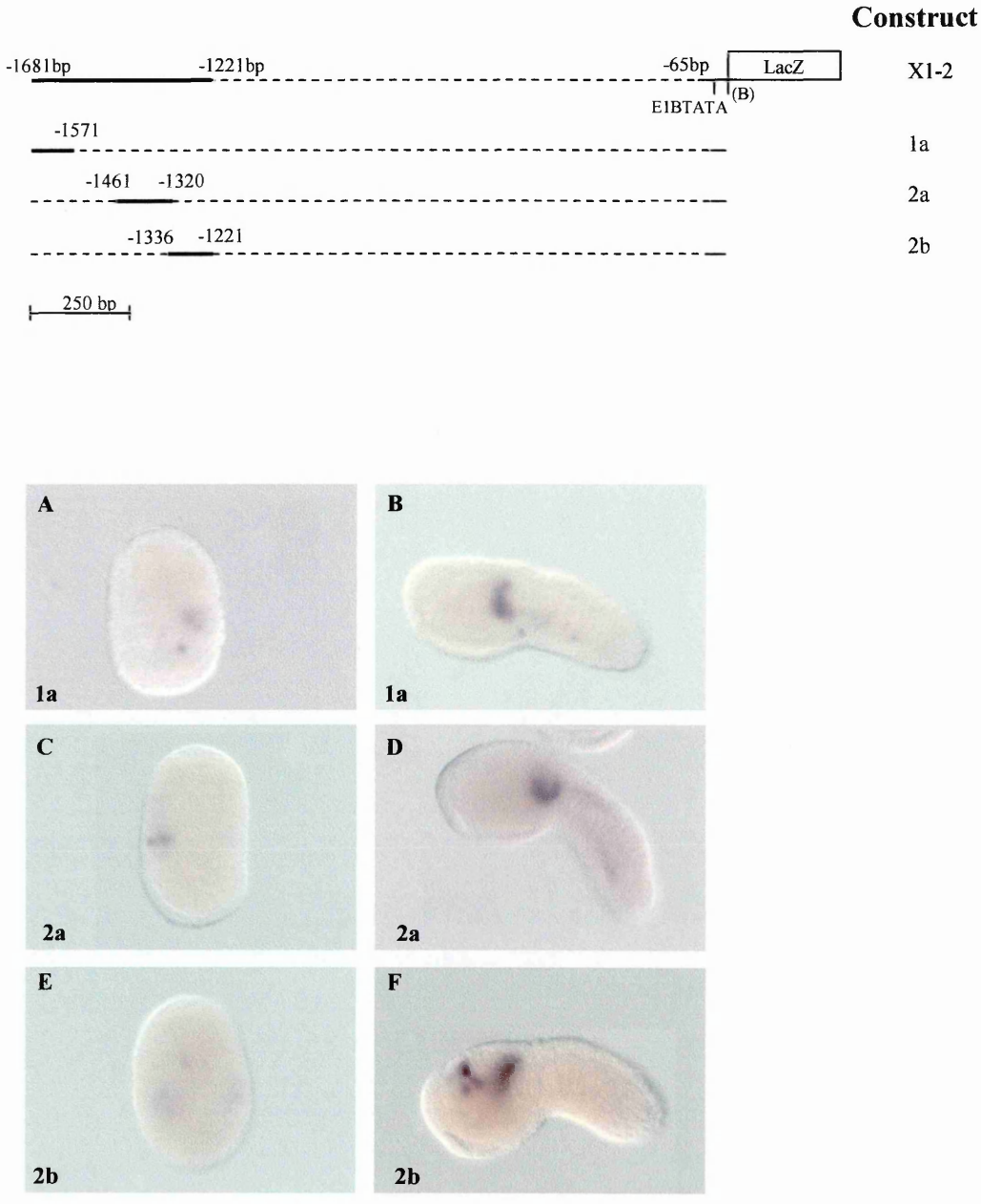


Figure 41

Expression of the 1a, 2a and 2b constructs in electroporated *Ciona* embryos by *LacZ RNA in situ* hybridization

Schematic representation of the constructs on the top; electroporated embryos on the bottom. **A,C.,E)** Dorsal view of neurula stage embryos and **B, D, F)** lateral view of tailbud stage embryos electroporated with constructs 1a, 2a, 2b respectively. *LacZ RNA* staining is present in both stages, ectopically in mesenchymal cells. In **F** is visible a signal also in the anterior part of the nervous system.

In contrast introduction into the embryos of construct 1b, which contains the promoter sequence extending from -1581 bp to -1451 bp, was able to restore staining in anterior territories (**fig. 42**); however the signal intensity resulted weaker compared to that obtained using the control plasmid Δ700 and the fragment X1-2 (**fig. 31** and **fig. 39**).

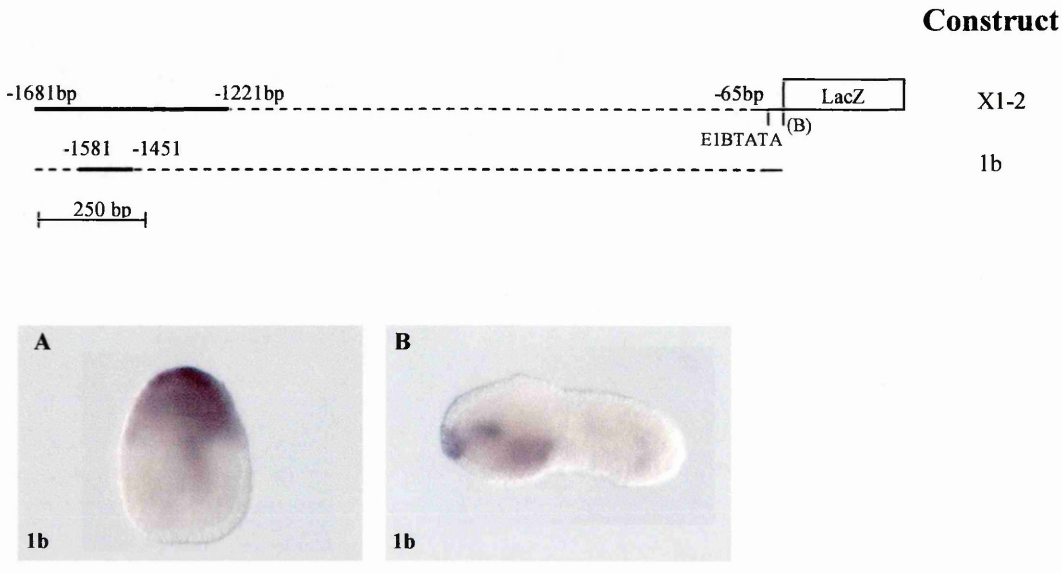


Figure 42

Expression of the 1b constructs in electroporated *Ciona* embryos by *LacZ RNA in situ* hybridization

Schematic representation of the constructs on the top; electroporated embryos on the bottom. **A)** Dorsal view of a neurula stage embryo. *LacZ RNA* staining is visible in the anterior part of the embryo including endoderm. **B)** Lateral view of a tailbud stage embryo. *LacZ RNA* staining is visible in epidermis, endoderm and anterior part of nervous system.

Moreover, about 20% of the embryos presented only an ectopic expression in mesenchyme cells

The number of the embryos, showing LacZ expression in the anterior region, and the staining intensity increased progressively with the introduction of the combined constructs 1a1b2a, 1b2a, 1b2a2b, compared to the single fragment 1b (**table 3**).

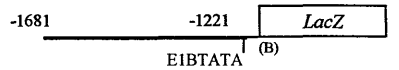
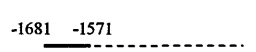
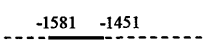
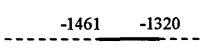
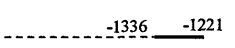
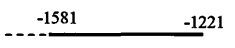
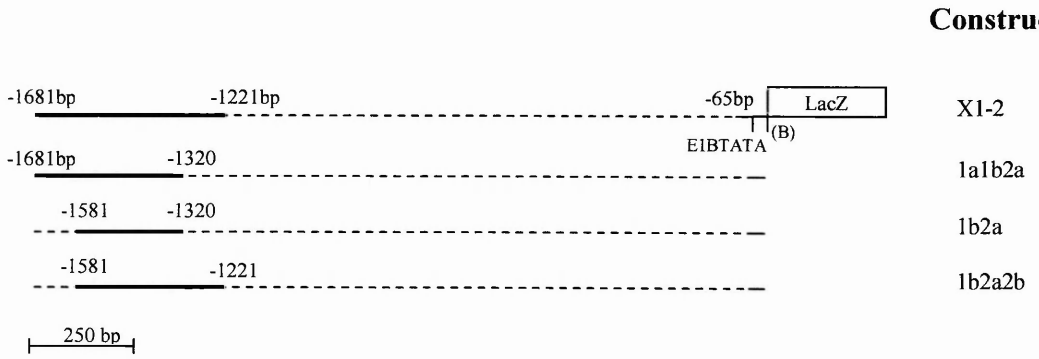
Construct name		Expression territory		Positives
		endoderm	ectopic mesenchyme	%
X1-2		+	+/-*	80
1a		-	+	30
1b		+	+/-*	50
2a		-	+	25
2b		-	+	30
1a1b2a		+	+/-*	70
1b2a		+	+/-*	60
1b2a2b		+	+/-*	85

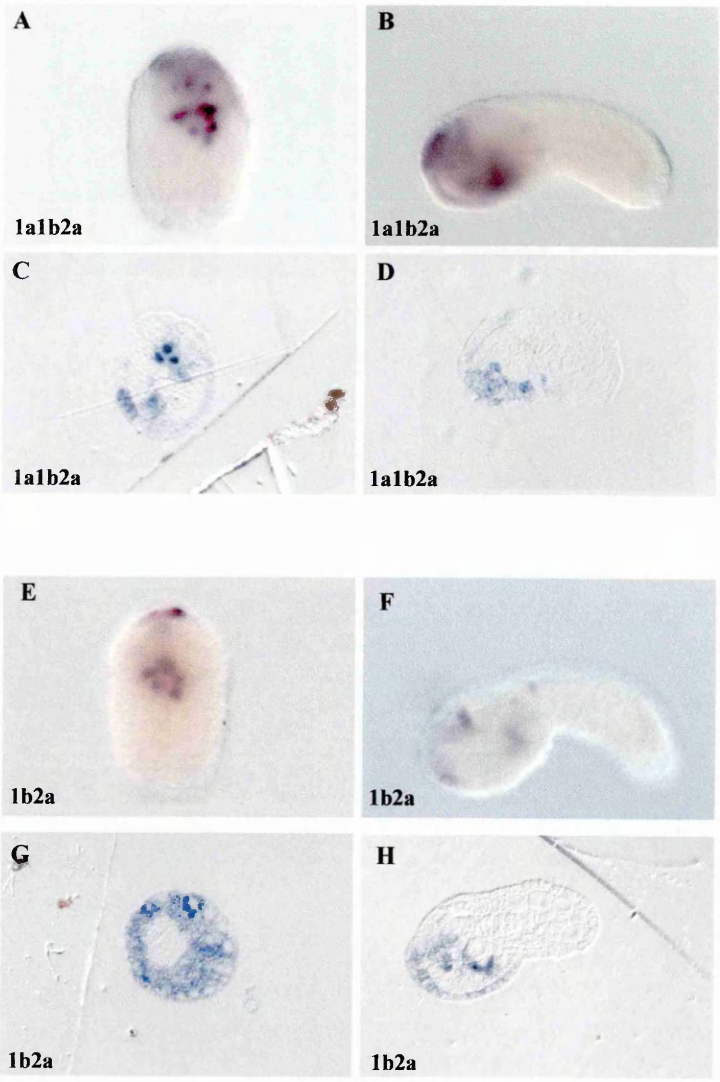
Table 3

Summary for the deletion analysis of X1-2 endoderm-specific promoter sequence. Schematic representation of promoter fragments analyzed in the electroporation experiments. The name of each construct is indicated on the left, the territory of reporter gene expression in the embryo and the percentage of stained embryos are shown on the right. The asterisk indicates that with the construct 1a1b2a, 1b2a2b and 1b2a some embryos (5–10%), besides the signal in the endoderm, present ectopic staining in the mesenchyme

Interestingly, the electroporation of these transgenes (1a1b2a, 1b2a2b, 1b2a) resulted in the development of embryos in which *LacZ* mRNA staining appeared localized anteriorly (**fig. 43**) more evident at the neurula stage (**fig. 43 A, E** and **I**) than at the tailbud stage (**fig. 43 B, F** and **L**). Endoderm expression was confirmed by semi-thin transverse sections, at the level of the anterior part of the neurula stage embryos (**fig. 43 C** and **G**) electroporated with 1a1b2a and 1b2a constructs and by semi-thin longitudinal sections, through anterior-posterior axis, of neurula embryos electroporated with 1b2a2b (**fig. 43 M**) transgene and of tailbud embryos electroporated with the three

constructs **(fig. 43 D, H** and **N)**. Endoderm expression was confirmed by semi-thin transverse sections, at the level of the anterior part of the neurula stage (**fig. 43 C** and **G)** electroporated with 1a1b2a and 1b2a constructs.





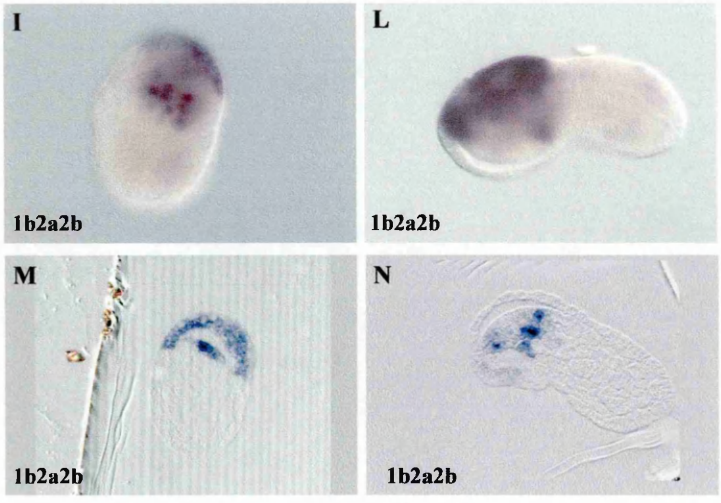


Figure 43

Expression of the 1a1b2a, 1b2a2b, 1b2a constructs in electroporated *Ciona* embryos by *LacZ RNA in situ* hybridization

Schematic representation of the constructs on the top; electroporated embryos on the bottom. **A**), **E**) and **I**) Dorsal view of a neurula stage embryo electroporated respectively with 1a1b2a; 1b2a; 1b2a2b constructs. *LacZ RNA* staining is present specifically in endoderm. **B**), **F**) and **L**) Lateral view of a tailbud stage embryo. **C**) and **G**) semi-thin transverse section at the level of the anterior part of neurula stage embryos electroporated respectively with 1a1b2a; 1b2a constructs. *LacZ RNA* staining is visible specifically in endoderm cells. **D**), **H**) and **N**) semi-thin longitudinal section through anterior-posterior axis of tailbud stage embryos electroporated respectively with 1a1b2a; 1b2a; 1b2a2b constructs. **M**) semi-thin longitudinal section through anterior posterior axis of neurula stage embryo electroporated with 1b2a2b construct shows anteriorly endoderm expression.

Based on these results, the elements controlling the expression in the endoderm appear to reside in the region extending from position -1581 to -1451 present in the 1b construct while the regions 2a, 2b and 1a seem to contain elements that cooperate for a full expression in the anterior region.

3.2 Genomic comparison: *Ciona intestinalis* vs *Ciona savignyi*

Identification of putative transmodules binding sites

The *C. savignyi* genome database, recently released (http://www-genome.wi.mit.edu/annotation/Ciona), represents a very useful tool to compare sequences of two phylogenetically and closely related species in order to find conserved regulatory sequences that could be functionally relevant.

The mVISTA comparative genomics tool (Mayor *et al.*, 2000) was used to carry out a sequence comparison between *C. intestinalis* and *C. savignyi sfrp* genes. Two FastA format sequences corresponding to almost 6 kb located around the *sfrp* coding region (3 kb upstream and 3 kb downstream) for each species were submitted to the Web server and the Vista.pdf file, showed in **figure 44**, is the visual representation of the alignment.

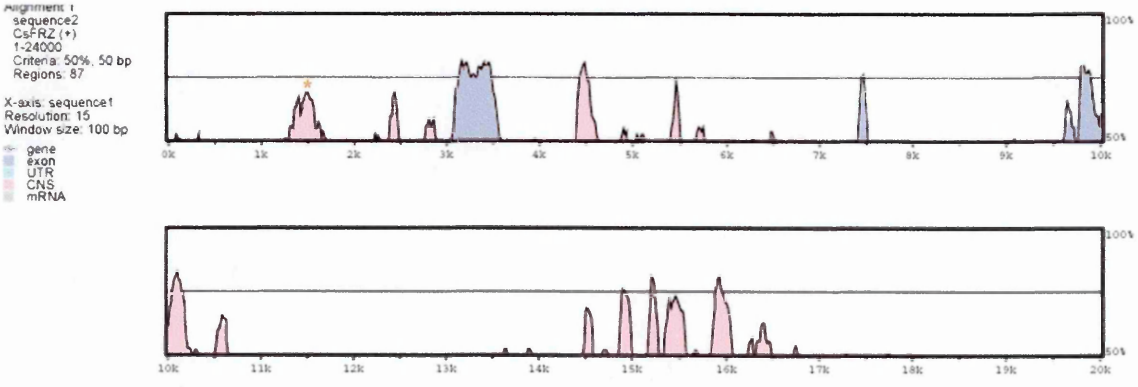


Figure 44

Sequence comparison of *C. intestinalis* and *C. savignyi* sfrp1/5 genes

Visual representation of the alignement and conserved regions found with the mVista comparative genomics tool. The dark blue regions indicated the exons, the pink regions the conserved non coding sequences. The peak indicated by the orange asterisk is in the region extending from position -1681 to -1221, the region of the X1-2 construct, containing the putative enhancer (s) for the expression in the anterior region of the embryo. The bases number are reported on the X axis and the percent of identity on the Y axis.

The results were in agreement with the *in vivo* analysis I carried out. In fact, I observed that the regulatory element(s) driving *Cisfrp1/5* expression in the anterior region

appeared to reside in the genomic fragment extending from position -1681bp to -1221bp from starting ATG codon.

To have an indication of the possible candidates able to bind the region extending from -1681bp to -1221bp, I submitted this sequence to the TRANSFAC database to check for possible binding sites recognised by known transcription factors. From the *in silico* analysis, different transcription factors appeared to recognise the sequence and the most conserved nucleotide positions are the core sequences for HNF/FOX (Forkhead) gene (Weigal and Jackle, 1990) and LIM-homeobox gene (Dawid *et al.*, 1995) and Gata zinc-finger factor families (Molkentin, 2000) (**fig. 45**).

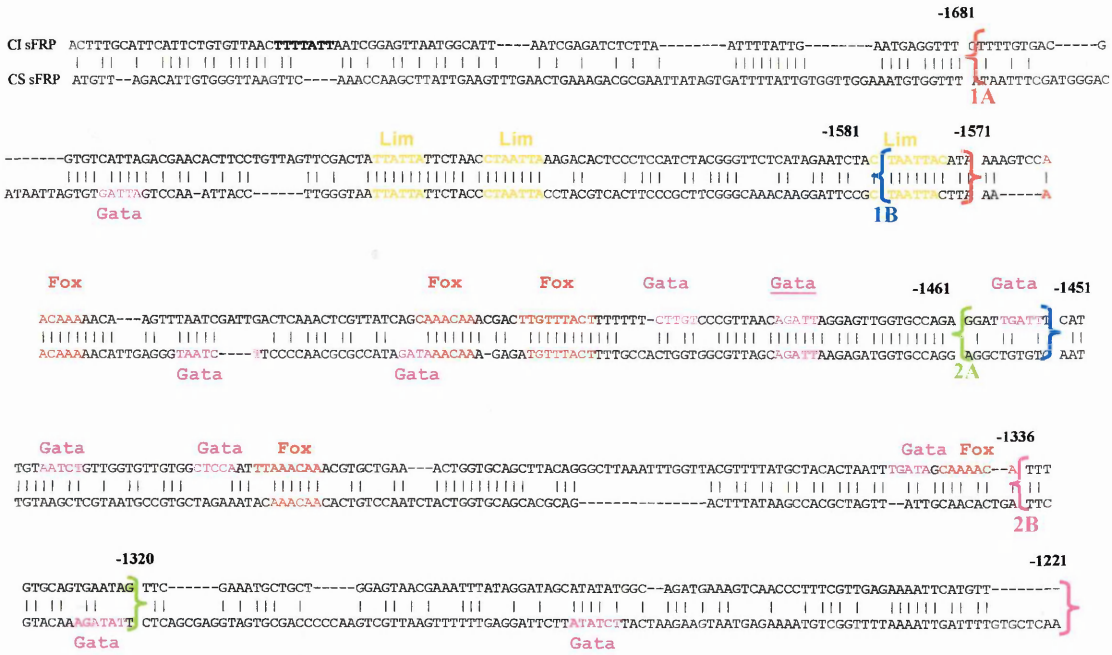


Figure 45

Sequence alignment between *C. intestinalis* X1-2 region and the *C. savignyi* corresponding fragment

CI sFRP corresponds to the *C. intestinalis* X1-2 fragment, the CS sFRP to the *C. savignyi* corresponding fragment. The core sequences of GATA, FOX and LIM are indicated in pink, red and yellow, respectively. 1a, 1b, 2a and 2b genomic fragments are indicated by brackets red, bleu, green and pink respectively.

3.3 *In vivo* analysis of the interaction between *Cititf1* and *CisFrp1/5* promoter region.

In order to verify the hypothesis of a relation between *Cititf1* and *CisFrp1/5* activation, I monitored the expression of *CisFrp1/5* gene, at the tailbud stage, in embryos electroporated with *CiBra/Cititf1* construct. The aim was to check if ectopic expression of *Cititf1* in the notochord, through *CiBra* promoter, could, in turn, induce an ectopic expression of *CisFrp1/5*. In **figure 46**, a control embryo, at the tailbud stage, electroporated with *CiBra/LacZ* construct shows *CisFrp1/5* expression in the anterior territories, including endoderm (**A**). In the embryo electroporated with *CiBra/Cititf1* construct (**fig. 46 B**), the territories of *CisFrp1/5* gene expression seem to be expanded compared to the control embryos (**A**). The aberrant development of the embryo in (**B**) is a typical phenotype due to the ectopic expression of *Cititf1* in the notochord (Spagnuolo and Di Lauro, 2002).

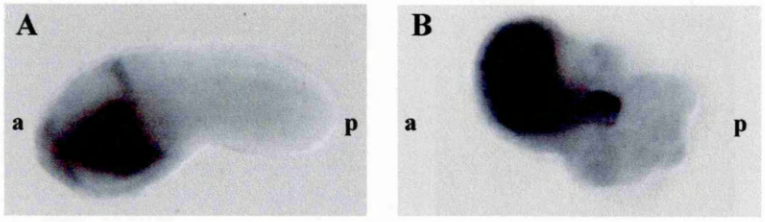


Figure 46

***In situ* hybridisation with *CisFrp1/5* riboprobe on *Ciona* embryos electroporated with *CiBra/Cititf1*.**

Lateral view embryo at tailbud stage (**A**) electroporated with *CiBra/LacZ* as control. Dorsal view embryo at taibud stage electroporated with *CiBra/Cititf1* and *CiBra/LacZ* constructs (**B**). a: anterior; p: posterior.

In a second series of experiments, I co-electroporated the construct *CiBra/Cititf1* and the construct 4/7E1b, containing the 1616 bases of *CisFrp1/5* promoter region upstream from E1bTATA and *LacZ* reporter gene. *CiBra* promoter drives the expression of *Cititf1* in the notochord, where normally it is not present. The hypothesis

was that in the case *Cititf1* could transactivate 4/7E1b promoter region, *LacZ* protein staining should be present ectopically in the notochord. **Figure 47**, shows embryos, at the middle tailbud stage, electroporated with the construct *4/7E1b/LacZ* alone (**A**), or together with *CiBra/Cititf1* (**B** and **C**). *LacZ* protein, in A, was present in the anterior territories, while in B and C the staining was broader and in some cases extended along the tail.

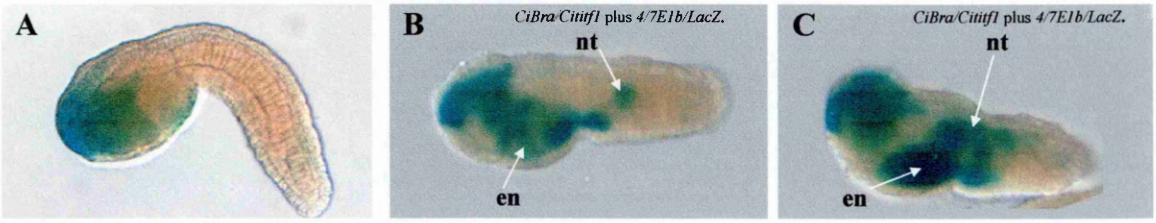


Figure 47

Expression of the *4/7E1bLacZ* construct in *Ciona* embryos co-electroporated with *CiBra/Cititf1*

A) Lateral view of control embryos at middle tailbud stage electroporated with *4/7E1bLacZ* at middle tailbud stage. *LacZ* staining is present in the anterior part of the embryo including endoderm. **B**) and **C**) Lateral view of an embryo co-electroprated with the two transgene, *CiBra/Cititf1* plus *4/7E1bLacZ*. *LacZ* staining is is broader and extended along the tail. These embryos (B and C) show also an aberrant phenotype due to *Cititf1* gene overexpression. en, endoderm; nt, notochord.

These results indicated that the ectopic expression of *Cititf1* gene in the notochord could induce an ectopic activation of *LacZ* in this tissue.

4 Analysis of the *Cititf1* minimal promoter S2

Previous studies on the transcriptional regulation of *Cititf1* have brought to the identification of a unique sequence of 27 nucleotides (S2), in *Cititf1* promoter region (Fanelli *et al.*, 2003), able to drive the expression of *LacZ* reporter gene specifically in the endoderm. Five copies head to tail of S2 oligonucleotide (5XS2) cloned upstream from the E1bTATA (5xS2E1b), are sufficient to drive a very strong expression of the

reporter gene both in the endoderm and ectopically in notochord and mesechyme (Lania *et al.*, data not published).

I used this promoter (5xS2E1b) to prepare a series of constructs that I tested in *Ciona*. Firstly, I attached *Cititf1* coding sequence to 5XS2E1b and checked the stage when *Cititf1* was activated by *in situ* hybridization experiment using *Cititf1* as probe. The resulting fusion construct, 5XS2E1b/*Cititf1*, was introduced in the fertilized eggs via electroporation; embryos were then collected and fixed at 32 cell, 64 cell, 110 cell, gastrula and neurula stages. The aim was to test, through *in situ* hybridization experiments using *Cititf1* as probe, when and where the gene was activated in notochord and endoderm lineage blastomeres.

The signal started, in most embryos, at the 32 cell stage in notochord blastomeres (compare **fig. 48 A** to **B, C, D**).

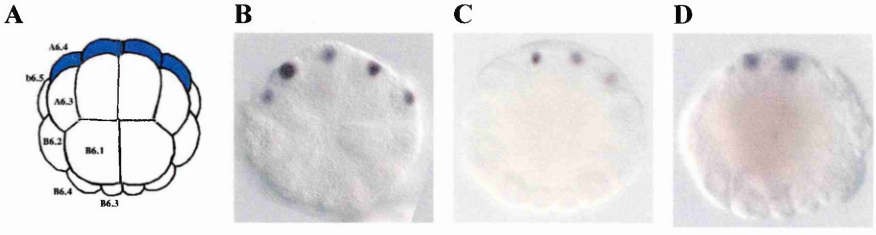


Figure 48

***Cititif1 in situ* hybridization on 32 cell stage embryos electroporated with 5XS2E1b/*Cititf1* construct**

A) Schematic representation of vegetal view of 32 cell stage. **B)**, **C)** and **D)** vegetal view of 32 cell stage embryos which show an ectopic *Cititif1* signal in notochord cells marked in blue in A).

At the 64 cell stage the expression domains of *Cititf1*, in most embryos, were localized into endodermal and notochord territories (compare **fig. 49 F** and **G**).

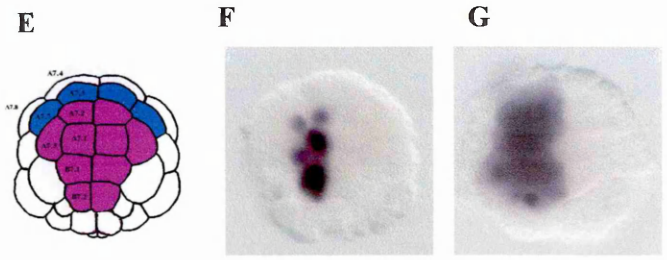


Figure 49

***Cititif1 in situ* hybridization on 64 cell stage embryos electroporated with 5XS2E1b/*Cititf1* construct**

E) Schematic representation of a vegetal view embryo at the 64 cell stage. **F**) and **G**) vegetal view of 64 cell stage embryos which show an ectopic *Cititif1* signal in notochord cells marked in blue in A) and in endodermal cells marked in violet in A).

At the 110 cell (**fig. 50 B, C, D**) and at gastrula stages (**fig. 51 F, G**) the signal was very strong both in the endoderm and notochord blastomeres.

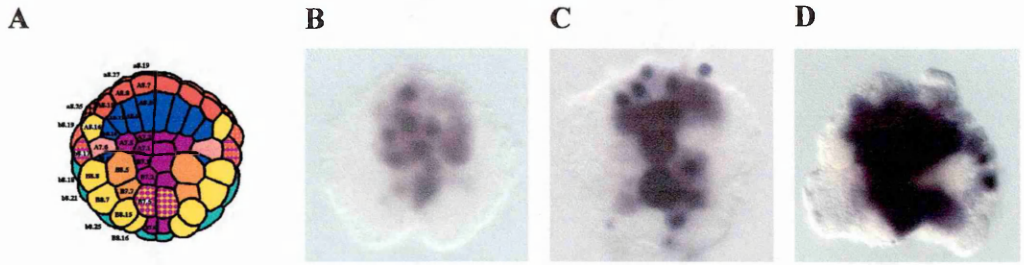


Figure 50

***Cititif1 in situ* hybridization on 110 cell stage embryos electroporated with 5XS2/*Cititf1* construct**

A) Schematic representation of endodermal precursor blastomeres (shown in violet) and notochord precursors of primary lineage (shown in blue) at the 110-cell stage. **B**), **C**) and **D**) vegetal view of 110 cell stage embryos which show *Cititif1* expression in the endodermal and notochord precursors.

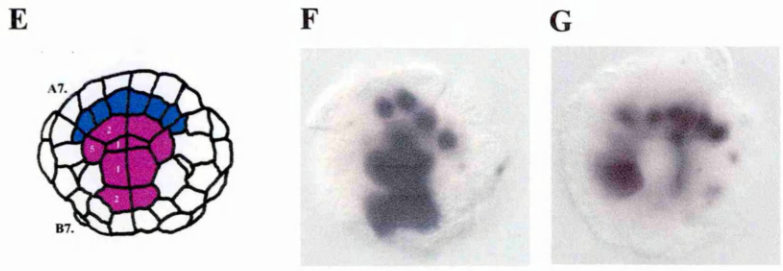


Figure 51

***Cititif1 in situ* hybridization on gastrula stage embryos electroporated with 5XS2E1b/*Cititf1* construct**

E) Schematic representation of endodermal precursor blastomeres (shown in violet) and notochord precursors of primary lineage (shown in blue) at the gastrula stage. **F**) and **G**) vegetal view of gastrula stage embryos which show *Cititif1* expression in the endodermal and notochord precursors precursors.

At these stages the embryos showed also a mutant phenotype, due to the ectopic expression of *Cititf1* in notochord blastomeres.

At neurula stage *Cititf1* mRNA persisted in these territories (**fig. 52 I** and **L** compared to **H**).

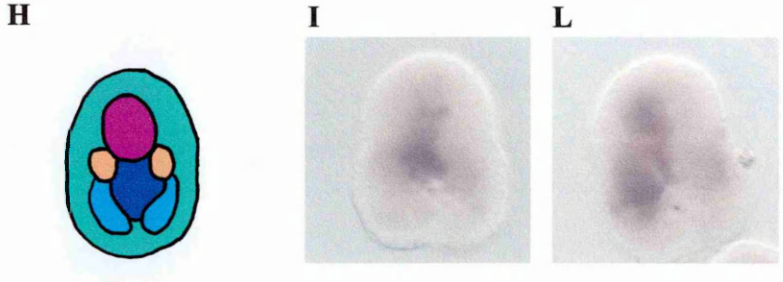


Figure 52

***Cititif1 in situ* hybridization on neurula stage embryos electroporated with 5XS2E1b/*Cititf1* construct**

H) Schematic representation of endodermal precursor blastomeres (shown in violet) at the neurula stage. **I**) and **L**) Dorsal view of neurula stage embryos which show *Cititif1* expression in the endodermal precursors. *Cititif1* expression domain is expanded into the blastomeres of notochord lineage shown in blue in **H**).

Therefore 5XS2E1b minimal promoter represents a useful tool to drive the expression of the genes cloned downstream from it, both in notochord and endodermal territories, since the 32 cell stage.

5 *CisFrp 1/5* expression is interfered by *Cititif1* expression

In a third set of experiments I analyzed *CisFrp1/5* mRNA in embryos in which *Cititf1* expression was interfered by using engrailed repressor domain. For this purpose the fusion construct *EnCiHD* (containing *Cititf1* homeodomain fused downstream from the Engrailed repressor domain) (Spagnuolo and Di Lauro, 2002) was cloned downstream from 5XS2E1b to create *5XS2E1b/EnCiHD*. The chimeric protein encoded by *EnCiHD* mRNA, expressed in the endodermal territories through the *5XS2E1b* promoter, should interfere with the function of the endogenous Cititf1 protein. As control I used the construct *5XS2E1b/EnCiHD-WF* that contains a double mutation in *Cititf1* homeodomain coding sequence at the residues 48-49 (WF/AA) already demonstrated to be crucial for the binding to the target DNA (Damante and Di Lauro, 1991, Spagnuolo and Di Lauro, 2002). These constructs were introduced into *Ciona* fertilized eggs via electroporation. The embryos were cultured up to the neurula and tailbud stages and then fixed for *in situ* hybridization experiments using *CisFrp1/5* mRNA as probe. In embryos electroporated with the construct *5XS2E1b/EnCiHD* both at neurula and tailbud stages (**fig. 53 B** and **E**) *CisFrp 1/5* expression appeared very faint (almost absent) in the anterior part of the embryo (in more than 50% of the embryos examined in 3 independent experiments) compared to the non electroporated control embryos (**fig. 53 A** and **D**) and to the control embryos electroporated with *5XS2E1b/EnCiHD-WF* (**fig. 53 C** and **F**).

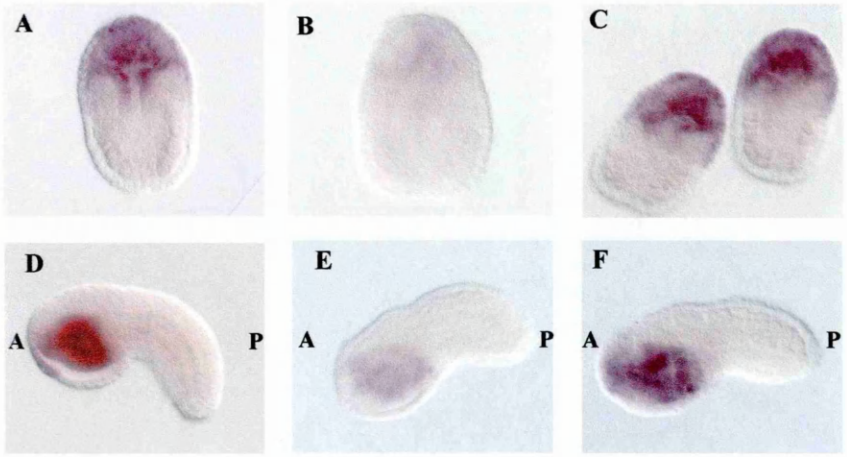


Figure 53

***CisFrp1/5 In situ* hybridization on embryos electroporated with *5XS2E1b/EnCiHD* and *5XS2E1b/EnCiHDWF* transgene.**

A) Dorsal view of neurula stage embryos wild type, **B)** electroporated with *5XS2E1b/EnCiHD* and **C)** electroporated with *5XS2E1b/EnCiHDWF*. **D)** Lateral view of tailbud stage embryos, wild type **E)** and electroporated with *5XS2E1b/EnCiHD*, **F)** electroporated with *5XS2E1b/EnCiHDWF*. The intensity of *CisFrp1/5* staining in B and E is very weak compared to that present in the wild type embryos (A and D) and in the embryos electroporated with *5XS2E1b/EnCiHDWF* (C and F). A: anterior; P: posterior.

This very weak signal was more evident at the neurula stage (**fig. 53 A, B, C**) than at the tailbud stage (**fig. 53 D, E, F**). These results suggest that interference with *Cititf1* function could in turn influence the *CisFrp1/5* expression during *Ciona* embryogenesis.

6 *In vivo* analysis of *CisFrp1/5* in development of *Ciona* embryos

The characteristic distribution of *CisFrp1/5* mRNA and its promoter activity suggest an involvement of this gene in the organization of endoderm. In an attempt to explore its putative function during *Ciona* development, I performed a series of experiments by driving *CisFrp1/5* ectopic expression, in notochord and mesenchyme,

under *5XS2E1b* promoter (*5XS2E1b/sFrp1/5* construct). The embryos electroporated with *5XS2E1b/sFrp1/5* construct were grown at 18°C and scored for normal or visible altered embryos at the neurula, tailbud and swimming larval stages. In each experiment the results were analyzed only when the control, non-electroporated embryos, showed at least 80% of normal development. No evident anomalies were detected both at neurula and at tailbud stages, while the effects appeared at the late tailbud stage and became evident at the larval stage. As shown in **figure 54**, by comparing with a non electroporated control (**fig. 54 A**), it was possible to identify, graded phenotypes with alterations in trunk-tail development. Some embryos presented a reduced (**fig. 54 B, D** and **E**) or a severely reduced tail (**fig. 54 C** and **F**).

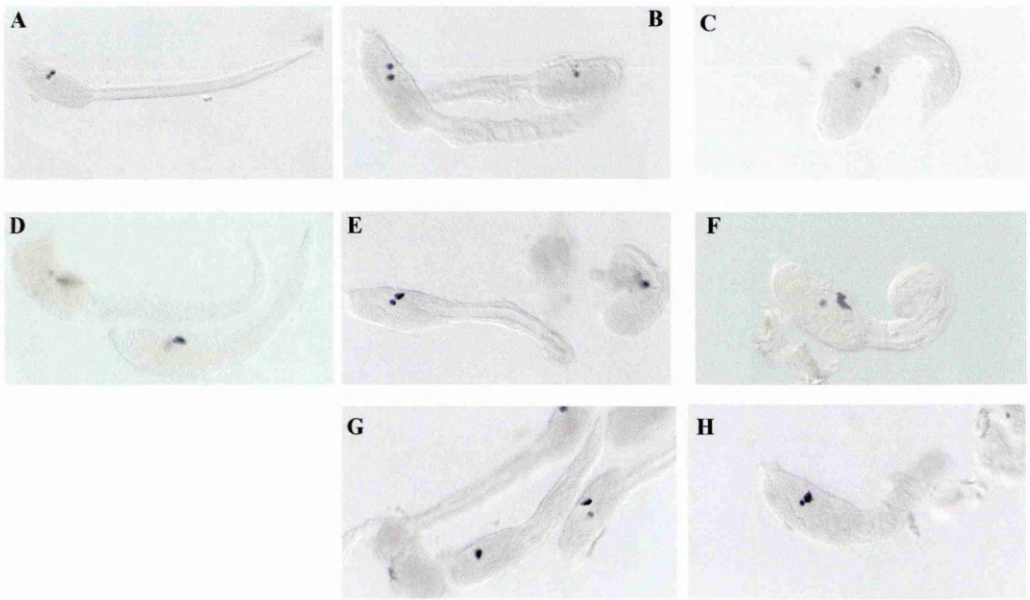


Figure 54

Overexpression of *CisFrp1/5*

A) Control non electroporated larva. **B-F**), larvae electroporated with 5X*S2E1b/sFrp1/5* transgene. Some representative embryos showed an altered phenotype, with altered trunk development, and a shorter, sometimes bent tail missing an organized notochord structure. The embryos in **G**) and **H**) which are electroporated with 5X*S2E1b/Cititf1* transgene, showed an altered phenotype which resembles that obtained with 5X*S2E1b/sFrp1/5* transgene electropration.

The anterior trunk structures, including the dorsal brain and the pigment cells, appeared normal (**fig. 54 B** and **E**) or rather enlarged (**fig. 54 D**) and sometimes extruded (**fig. 54 C** and **F**). This phenotype, in several respects, is comparable with that obtained by *Cititf1* overexpression using the constructs 5XS2E1b/*Cititf1*. (**fig. 54 G** and **H**) particularly in relation to tail structure and head-trunk organization.

However, as represented in the **table 4**, the percentage of phenotipically altered *Cititf1* overexpressing larvae was higher than the percentage of phenotipically altered *CisFrp1/5* overexpressing larvae.

		Larva stage	
	Construct name	Normal/Total	%
Exp1	Control (no electroporated)	85/87	98
	5XS2E1b/CisFrp1/5	16/36	44
	5XS2E1b/Cititf1	12/37	32
Exp2	Control (no electroporated)	65/72	90
	5XS2E1b/CisFrp1/5	16/33	48
	5XS2E1b/Cititf1	8/30	26
Exp3	Control (no electroporated)	110/137	80
	5XS2E1b/CisFrp1/5	13/21	62
	5XS2E1b/Cititf1	45/57	79

Table 4

Effects of *Cittf1* and *sFrp1/5* overexpression on development of *Ciona* embryos

In three separate experiments (1-3) the numbers of normal or visibly altered embryos were scored at larval stage, and the percentage of normal embryos is indicated on the right. The constructs used in electroporation experiments are indicated on the left.

Discussion

Given the knowledge about embryo development and the possibility to rapidly identify cis-regulatory DNA sequences by using electroporation of chimeric reporter genes (Corbo *et al.*, 1997a; Di Gregorio, 2002; Takahashi *et al.*, 1999), the ascidian *Ciona intestinalis* is an ideal experimental model to explore the molecular mechanisms underlying the expression and function of genes involved in development.

The aim of this thesis was the identification of putative targets of *Cititf1* that could be involved in late differentiation of the endoderm. This germ layer, scarsely studied in the past, has been recently the subject of different studies that have started to characterize the molecular determinants responsible for endoderm early differentiation. This is mostly established during gastrulation (Shivdasani, 2002) and results into a gradual regionalization of this tissue into anteroposteriorly and dorsolaterally divided organs (Grapin-Botton, 2000). Compared with the study of early endoderm differentiation and late organogenesis, however, the molecular pathways that control this process remain largely unknown. In *Ciona*, the larval endodermal cells are not differentiated and only after metamorphosis they give rise to a variety of adult endodermal organs, including digestive tract, endostyle and branchial sac. Although at the larval stage no signs of organogenesis are evident, it has been demonstrated that developmental fates of larval endodermal cells after metamorphosis are almost invariant, indicating that a certain degree of regionalization already exists in this tissue (Hirano and Nishida 2000). *Cititf1* is a clear confirmation of this projection since its transcript is present at the larval stage in the anterior-ventral part of head endoderm, a region that, after metamorphosis, gives rise to the endostyle (the thyroid ancestor) where *Cititf1* is still expressed.

Besides playing a role in late embryogenesis, *Cititf1* is also a molecular determinant during early endoderm differentiation (Ristoratore *et al.*, 1999). Given this information, I considered that *Cititf1* could represent a key factor to identify genes downstream from it and eventually involved in early regionalization of endodermal territories from neurula stage.

Through a subtractive hybridization screen (SSH) between *Cititif1*-over expressing embryos and control embryos, I have identified *CisFrp1/5* as a potential target of *Cititif1*.

The SSH, a technique largely used to isolate genes differentially expressed, suffers, however, from some disadvantages that can invalidate its potentialities. In my case the experiments suggested by the manufacturers to check the whole procedure, indicated that some of the clones isolated in this screening were not differentially expressed upon *Cititf1* overexpression (data not shown). Moreover, the percentage of endoderm-expressed genes (10/200) recovered during this analysis was much lower than expected. It is possible that the subtracted sample contained some cDNAs that corresponded to mRNAs common to both the tester and control samples. This background problem occurs when very few mRNA species are differentially expressed in tester and control samples. My subtraction was done between *Cititf1* over expressing embryos producing, potentially, high amounts of the transcripts of interest, and control embryos, containing normal levels of the same RNAs. One possibility is that the two RNA populations contained a limited set of differentially expressed transcripts and low quantitative differences that could have been minimized further by a possible overcycling, during the initial PCR amplification steps, causing the low efficiency of the subtraction.

Besides these technical problems, however, this approach led to the identification of *CisFrp1/5* gene as a potential target of *Cititf1*. *CisFrp1/5* belongs to a recently discovered family, the secreted frizzled related proteins, involved in the

regulation of Wnt signal, a molecular cascade exerting so diverse function during embryonic differentiation.

Until the sequencing of the *Ciona* genome, sFRPs were considered to be exclusive to vertebrate taxa (with the exception of the sea urchin *Strongylocentrotus purpuratus* (Illies *et al.*, 2002). The release of the draft sequence of *C. intestinalis* genome (Dehal *et al.*, 2002) identified four sFRP genes in this chordate ancestor, *CisFRP1/5*, *CisFRP2*, *CisFRP3/4-a* and *CisFRP3/4-b*. The isolation of *CisFrp1/5* during my screen prompted me to analyze this gene in *Ciona*, hoping that this simple, but evolutionary strategic experimental model, could shed light on the pathways in which these proteins are involved.

By whole mount *in situ* hybridization, *CisFrp1/5* transcript was identified in CNS, mesenchyme and notochord blastomere precursors at the 110 cell stage; later, at the neurula stage, the signal was localized in the most anterior part of the embryos, which includes the ectoderm and the endoderm (**fig. 22**). The endoderm, at this stage, is completely covered by the ectoderm: semi-thin longitudinal and transverse sections were therefore used to reveal the presence of *CisFrp1/5* signal in the endoderm, besides the external ectoderm. The expression territories were conserved up to the tailbud stage. At the larval stage the transcript was localized in the brain vesicle and in the endodermal cells ventrally to the brain vesicle (**fig. 23**). The onset of *CisFrp1/5* expression in the endoderm, from the neurula stage, identified it as a potential target of *Cititf1*.

Additional support for this possibility was obtained by analyzing *CisFrp1/5* in mutant *Ciona* tadpoles that contain excess endoderm tissue due to the misexpression of *Cititf1*. In these embryos the territories of *CisFrp1/5* expression were expanded, when compared to the control embryos (**fig. 46**).

On the other hand, interference with *Cititf1* function by using engrailed repressor domain, appeared to lower the expression of *CisFrp 1/5* (**fig. 53**). These data are indicative of a link between *Cititf1* activation and *CisFrp 1/5* - related expression. Furthermore, misexpression of the gene in notochord and mesenchyme, has similar effects to that of the *Cititf1* misexpression phenotype (**fig. 54**), but with a lower penetrance (**table 4**). This result can be interpreted in different ways. It is known that *Citif1* is expressed in a topical moment, namely at 64/110 cells stage when the endodermal blastomeres are fate restricted. *CisFrp 1/5*, instead, appears in endodermal cells at the neurula stage, when the fate of endodermal cells is further restricted. On the other hand the differences could reside in the role assigned to these genes: *Cititf1*, as a molecular determinant of endoderm specification, *CisFrp 1/5* as a factor devoted to the completion of endodermal territories.

Regulation of the *CisFrp 1/5* gene

Since *CisFrp 1/5* was isolated during a screening for putative *Cititf1* targets, the study of its transcriptional regulation could therefore be helpful to identify modules controlling its tissue-specific activation in endodermal territories and its potential relation with *Cititf1*.

Until now, little information is available about the *cis*-regulatory elements that control *sFrps* expression in different organisms. Identification of minimal transcriptional elements within promoter regions is of striking interest, since they contribute to the clarification of the involvement of the genes that regulate specific genomic pathways. Moreover by a computational analysis, it is possible to identify in these elements, recognition sites for putative interacting factors. To start my analysis, a 5' genomic sequence was amplified by PCR reaction on *Ciona intestinalis* genomic DNA. This region extends between -1681bp and -65bp upstream the ATG starting

codon of the gene. The construct, containing the 1.6 kb fragment, was sufficient to reproduce the complete *CisFrp1/5* gene spatial expression pattern (**fig. 27** and **fig. 28**). Also, the onset of β-galactosidase expression at neurula stage was well conserved. In almost all the experiments I checked the expression of the transgene through *LacZ* RNA *in situ* hybridization, since the time necessary to accumulate detectable amounts of LacZ protein is often delayed, compared to the appearance of the endogenous gene.

Using computational tools, the 1.6 kb sequence was found to contain two recognition sites for *Cititif1* transcription factor. However, point mutations in these sites, either both or separately, did not cause evident loss in any aspect of the staining pattern, indicating that these sites are not involved in the reporter activation. It can be hypothesized that cooperation among transcription factors is needed for a full *CisFrp1/5* expression in the endoderm.

Using constructs truncated at the 3' (Δ700) or at the 5' end (Δ1000) I found that the Δ700 fragment (extending from -1681bp to -981bp) was able to reproduce *CisFrp1/5* expression in the anterior region of embryos, including endoderm, as assessed by sections, both at the neurula and tailbud stages (**fig. 31**). In contrast embryos transgenic for Δ1000 construct showed only an ectopic expression of *LacZ* mRNA in mesenchyme cells (**fig. 32**). This is not surprising, since the ability of specific promoter fragments to drive ectopic expression in the mesenchyme cells been already reported in studies on the transcriptional control of other ascidians genes (Corbo *et al.*, 1997b; Di Gregorio and Levine, 1999; Locascio *et al.*, 1999, Fanelli *et al.*, 2003).

The Δ700 region was further dissected up to the identification of the fragment X1-2, localized between -1681bp and -1221, able to control reporter activation in the anterior region, including the endoderm (**fig. 39**).

Using computational tools, I performed a phylogenetic foot print analysis and compared genomic regions from *C. intestinalis* and *C. savignyi*. This analysis allowed

me to obtain important information about the conservation of the regulatory regions. An interesting result from this analysis was the perfect match with the results obtained by *in vivo* experiments. In fact, the most *C. intestinalis* versus *C. savignyi* conserved region extends from the position -1681bp to -1221bp (**fig. 44**).

Serial of 5' truncations of X1-2, starting from the position -1681bp, permitted a more detailed analysis of this region, leading to the identification of the sequence extending between -1581bp to -1451bp (**fig. 42**), able to restore some staining in the anterior part of the embryos. However, the data indicated that additional sequences are required to achieve a stronger anterior expression as demonstrated by electroporations of the constructs 1a1b2a, 1b2a and 1b2a2b. In particular, removal of the sequence (-1320 bp to -1221 bp) localized at the most 3' end of the region -1681bp to -1221bp led to a decrease in the expression efficiency of the reporter (**fig. 41**), as demonstrated by the lower percentage of stained embryos compared with the control (**table 3**).

At the neurula stage, the anterior region of *Ciona* embryos includes the internal endoderm and the external ectoderm, which is differentiated into epidermis and neural plate (CNS precursor). The evidence I have collected so far indicates that the transcriptional regulation of *CisFrp1/5* expression in epidermis, CNS and endoderm is mediated by a compact region of the 360 bp sequence. Despite my efforts, I have not been able to separate the modules responsible for the expression in each specific territory. I can suggest that these modules are tightly interconnected and that this could be due to the putative functional features of *CisFrp1/5.* In fact, as already mentioned, sFRPs are modulators of Wnt genes that are potent morphogens and are involved in very different processes including gastrulation (Jones, 2002). In *C. intestinalis* ten *Wnt* genes are present and their function during embryogenesis has not been analyzed so far. The only solid evidence comes from a Wnt effector, β-catenin, which has been demonstrated to control early endoderm differentiation. The presence of *CisFrp1/5* in

the most anterior region could be related to a role played by this gene in establishing antero-posterior polarity, through a control of a specific Wnt factor acting during gastrulation or later. If this is the case, it is easy to argue that *CisFrp1/5* expression must to be finely and precisely controlled and this can be accomplished through close shared elements that influence its expression in all three tissues. Using computational tools, I analyzed the region between position -1681bp to -1221bp and found that it contains many interspersed LIM, Gata, TTF1 and HNF/FOX recognition sites **(fig. 55)**.

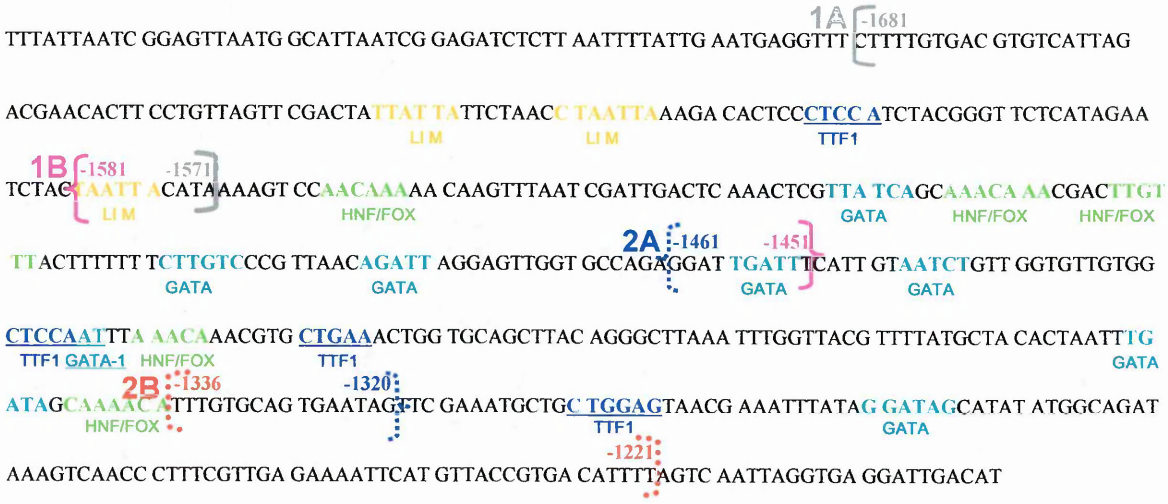


Figure 55

Result of TRANSFAC analysis

Genomic region from position -1681bp to -1221bp contained in the X1-2 construct presents recognition sites for LIM (indicated in yellow), GATA (indicated in dark green), TTF1 (indicated in bleu), and HNF/FOX (indicated in light green) factors. The genomic regions contained in the constructs 1a, 1b, 2a and 2b are indicated by grey, pink, blue and red brackets, respectively.

Gata, Lim and HNF/FOX factors are involved both in endoderm and CNS differentiation in diverse taxa (Nardelli *et al.*, 1999; Laverriere *et al.*, 1994; Neave *et al.*, 1995; Lantz and Kaestner, 2005; Hobert and Westphal, 2000).

Moreover, in *Ciona*, Gata a, has been already demonstrated to be involved in neural tissue differentiation (Bertrand *et al.*, 2003), while one of the HNF/FKH family

members, *Ci-fkh,* is expressed in notochord, endoderm, and the rudimentary floor plate of the CNS (Di Gregorio *et al.*, 2001). In *Ciona savignyi*, a Lim factor, *Cs-lhx3*, controls early endoderm differentiation (Satou *et al.*, 2001).

As the organization of the X1-2 enhancer with "endoderm and CNS factor recognition sites" tightly alternate and are linked together, this seems to confirm the hypothesis of a 'all or none' regulation of *CisFrp1/5* in the anterior part of the embryos. This structure, however, makes a precise dissection of the elements necessary for the activation of *CisFrp1/5* in each specific territory very difficult. To better define this region, mutations of each class of sites, either alone or in different combinations, could be tried. In another approach, it could be verified if members of Gata, TTF1 and HNF/Fkh are able to activate the reporter gene downstream from *CisFrp1/5* promoter.

I have already performed some preliminary test, by using *Cititf1* in co-electroporation experiments. This strategy consists in inducing expression of *Cititif1* in a territory where this factor is normally not expressed. The *CiBra*-promoter, driving the expression of the reporter genes in notochord (Corbo *et al.*, 1997b), was chosen to induce the expression of *Cititif1*. If *Cititif1* interacts with *CisFrp1/5* promoter (*4/7E1b*) present in co-electroporated *CiBra/Cititif1-4/7E1b/LacZ* embryos, *LacZ* reporter gene expression will be visible in the notochord. Actually LacZ protein, in co-electroporated embryos, was detected in broader territories, compared with the control embryos, and in some cases the staining was extended along the tail. These results, together with the promoter analysis, suggest that a cross-regulation may exist, between *Cititif1* and *CisFrp1/5*. In this respect, it will be really interesting to analyze the involvement of the TTF1 in *CisFrp1/5* endoderm-activation, by specific point mutation of the recognition sites identified in the X1-2 region.

Conclusions

The results I have obtained indicate that *CisFrp1/5* appears to be a factor that contributes to the regionalization of the whole embryo. Infact the expression of this gene demarcates the most anterior region of the embryo from the late gastrula/early neurula stages. Furthermore, *CisFrp1/5* may also contribute in the establishment of the boundaries of the endodermal territories. Indeed, it has to be noticed that both the endogenous gene and the reporter gene are expressed exclusively in the anteriormost endodermal blastomeres.

The precise and specific *CisFrp1/5* expression seems to be controlled by a compact module, X1-2, that is fundamental for its activation in the anterior region of *Ciona* neurula embryos; most probably accessory sequences can be required to cooperate for a full *CisFrp1/5* expression.

Although many questions remain still unanswered, the ascidian system has allowed me to reveal new potential role for these proteins during embryogenesis, and to start to shed light on their transcriptional regulation and on the pathways in which they are involved.

Bibliography

Aberle, H., Bauer, A., Stappert, J., Kispert, A. and Kemler, R. (1997). ß-catenin is a target for the ubiquitin-proteasome pathway. EMBO J; **16**: 3797-3804.

Afouda, BA., Ciau-Uitz, A. and Patient, R. (2005). GATA 4, 5 e 6 mediate TGFβ maintenance of endodermal gene expression in *Xenopus* embryos. Development; **132:** 763-774.

Alexander, J. and Stainier, DYR. (1999). A molecular pathway leading to endoderm formation in zebrafish. Curr Biol; **9:** 1147-1157.

Alonso, L. and Fuchs E. (2003). Stem cells of the skin epithelium. Proc Natl Acad Sci. USA; **100**:11830-5.

Andreé, B., Duprez, D., Vorbusch, B., Arnold, H. and Brand, T. (1998). BMP-2 induces ectopic expression of cardiac lineage markers and interferes with somite formation in chicken embryos. Mech Dev; **70:** 119-131.

Angerer, LM. and Angerer, RC. (2000). Animal-vegetal axis patterning mechanisms in the early sea urchin embryo. Dev Biol; **218**: 1-12.

Aros, B. and Viragh, S. (1969). Fine structure of the pharynx and endostyle of an ascidian (*Ciona intestinalis*). Acta Biol; **20**: 281-297.

Bafico, A., Gazit, A., Pramila, T., Finch, PW., Yaniv, A. and Aaronson, S. A. (1999). Interaction of frizzled related protein (FRP) with Wnt ligands and the frizzled receptor suggests alternative mechanisms for FRP inhibition of Wnt signaling. J Biol Chem; **274**: 16180-16187.

Bafico, A., Liu, G., Yaniv, A., Gazit, A. and Aaronson, S. A. (2001). Novel Antagonists of the Wnt signalling pathway mechanism of Wnt signalling inhibition mediated by Dickkopf-1 interaction with LRP6/Arrow. Nat Cell Biol; **3**: 683-686.

Baker, NE. (1987). Molecular cloning of sequences from wingless, a segment polarity gene in *Drosophila*: the spatial distribution of a transcript in embryos. EMBO J; **6**: 1765-1773.

Banyai, L. and Patthy, L. (1999). The NTR module: domains of netrins, secreted frizzled related proteins, and type I procollagen C-proteinase enhancer protein are homologous with tissue inhibitors of metalloproteases. Protein Sci; **8**: 1636-1642.

Baranski, M., Berdougo, E., Sandler, JS., Darnell, DK. and Burrus, LW. (2000). The dynamic expression pattern of frzb-1 suggests multiple roles in chick development. Dev Biol; **217**: 25–41.

Beddington, RS. and Robertson, EJ. (1999). Axis development and early asymmetry in mammals. Cell; **96**: 195-209.

Bertrand, V., Hudson, C., Caillol, D., Popovici, C. and Lemaire, P. (2003). Neural tissue in ascidian embryos is induced by FGF9/16/20, acting via a combination of maternal GATA and Ets transcription factors. Cell; **115**: 615-27.

Bhanot, P., Brink, M., Samos, CH., Hsieh, JC., Wang, Y., Macke, JP., Andrew, D., Nathans, J. and Nusse, R. (1996). A new member of the frizzled family from *Drosophila* functions as a Wingless receptor. Nature; **382**: 225–230.

Bienz, M. and Clevers, H. (2000). Linking colorectal cancer to Wnt signaling. Cell; **103**: 311-20.

Bouwmeester, T., Kim, S., Sasai, Y., Lu, B. and de Robertis, EM. (1996). Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer. Nature; **382**: 595-601.

Bowerman, B., Eaton, BA. and Priess, JR. (1992). skn-1, a maternally expressed gene required to specify the fate of ventral blastomeres in the early *C. elegans* embryo. Cell; **68:** 1061-1075.

Bradley, L., Sun, B, Collins-Racie, L., LaVallie, E., McCoy, J. and Sive, H. (2000). Different activities of the frizzled-related proteins frzb2 and sizzled2 during *Xenopus* anteroposterior patterning. Dev Biol; **227**: 118–132.

Bronner, G., Chu-LaGraff, Q., Doe, CQ., Cohen, B., Weigel, D., Taubert, H. and Jackle, H. (1994). Sp1/egr-like zinc-finger protein required for endoderm specification and germ-layer formation in *Drosophila*. Nature; **369**: 664-8.

Brott, BK. and Sokol, SY. (2002). Regulation of Wnt/LRP signaling by distinct domains of Dickkopf proteins. Mol Cell Biol; **22**: 6100-6110.

Cabrera, CV., Alonso, MC., Johnston, P., Phillips, RG. and Lawrence, PA. (1987). Phenocopies induced with antisense RNA identify the wingless gene. Cell; **50**: 659-63.

Cameron, CB., Garey, JR. and Swalla, BJ. (2000). Evolution of the chordate body plan: new insights from phylogenetic analyses of deuterostome phyla. Proc Natl Acad Sci USA; **97**: 4469-4474.

Capdevila, J., Tabin C. and Johnson, RL. (1998). Control of dorsoventral somite patterning by Wnt-1 and 3-catenin. Dev Biol; **193:** 182-194.

Chang, JT., Esumi, N., Moore, K., Li, Y., Zhang, S. et al. (1999). Cloning and characterization of a secreted frizzled-related protein that is expressed by the retinal pigment epithelium. Hum Mol Genet; **8**: 575-583.

Clements, D., Rex, M. and Woodland, HR. (2001). Initiation and early patterning of the endoderm. Int Rev Cytol; **203:** 383-446.

Conklin, EG. (1905). The organization and cell-lineage of the ascidian egg. Acad Nat Sci Phil; **13**: 1-119.

Conlon, FL., Lyons, KM., Takaesu, N., Barth, KS., Kispert, A., Hermann, B. and Robertson, EJ. (1994). A primary requirement for Nodal in the formation and maintenance of the primitive streak in the mouse. Development; **120:** 1919-1928.

Corbo, JC., Erives, A., Di Gregorio, A., Chang, A. and Levine, M. (1997a). Dorsoventral patterning of the vertebrate neural tube is conserved in a protochordate. Development; **124**: 2335-2344.

Corbo, J.C., Levine, M. and Zeller, RW. (1997b). Characterization of a notochord-specific enhancer from the *Brachyury* promoter region of the ascidian, *Ciona intestinalis*. Development; **124**: 589-602.

Corbo, JC., Fujiwara, S., Levine, M. and Di Gregorio, A. (1998). Suppressor of hairless activates brachyury expression in the *Ciona* embryo. Dev Biol; **203**: 358-368.

Dale, L., Howes, G., Prive, B. and Smith, J. (1992). Bone morphogenetic protein 4: A ventralizing factor in early *Xenopus* development. Development; **115:** 573-585.

Damante, G. and Di Lauro R. (1991). Several regions of the Antennapedia and thyroid transcription factor 1 homeodomains contribute to DNA binding specificity. Proc Cell Dev Biol; **13**: 611-667.

David, NB. and Rosa, FM. (2001). Cell autonomous commitment to an endodermal fate and behaviour by activation of Nodal signalling. Development; **128**: 3937–3947.

Davidson, EH., Rast, JP., Oliveri, P., Rasnick, A., Calestani, C. et al. (2002). A genomic regulatory network for development. Science; **295**: 1669-1678.

Dehal, P., Satou, Y., Campbell, RK., Chapman, J., Degnan, B. et al. (2002). The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. Science; **298**: 2157-2167.

Dennis, S., Aikawa, M., Szeto, W., D'Amore, PA. and Papkoff, J. (1999). A secreted frizzled related protein, FrzA, selectively associates withWnt-1 protein and regulates wnt-1 signaling. J Cell Sci; **112**: 3815–20.

Dawid, IB. Toyama, R. and Taira, M. (1995). LIM domain proteins. CRAcad Sci III; **318**: 295-306.

Devereux, J., Haeberli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res; **12**: 387-395.

Dickmeis, T., Mourrain, P., Saint-Etienne, L., Fischer, N., Aanstad, P. et al. (2001). A crucial component of the endoderm formation pathway, casanova, is encoded by a novel sox-related gene. Gene and Dev; **15**: 1487-1492.

Di Gregorio, A. and Levine, M. (1998). Ascidian embryogenesis and the origins of the cordate body plan. Curr Opin Genet Dev; **8**: 457-463.

Di Gregorio, A. and Levine, M. (1999). Regulation of Ci-Tropomyosin-like, a Brachyury target gene in tha ascidian, *Ciona intestinalis*. Development; **126**: 5599-5609.

Di Gregorio, A., Corbo, JC. and Levine, M. (2001). The regulation of forkhead/HNF-3β expression in the *Ciona* embryo. Dev Biol; **229**: 31-43.

Di Gregorio, A. and Levine, M. (2002). Analyzing gene regulation in ascidian embryos: new tools for new perspectives. Differentiation; **70**: 132-139.

Esteve, P., Morcillo, J. and Bovolenta, P. (2000). Early and dynamic expression of cSfrp1 during chick embryo development. Mech Dev; **97**: 217–221.

Fainsod, A., Steinbeisser, H. and De Robertis, EM. (1994). On the function of BMP-4 in patterning the marginal zone of the Xenopus embryo. EMBO J; **13:** 5015-5025.

Fan, CM. and Tessier-Lavigne, M. (1994). Patterning of mammalian somites by surface ectoderm and notochord: Evidence for sclerotome induction by a hedgehog homolog. Cell; **79:** 1175-1186.

Fan, CM., Porter JA., Chiang C., Chang DT., Beachy PA, and Tessier-Lavigne M. (1995). Long-range sclerotome induction by Sonic hedgehog: Direct role of the amino-terminal cleavage product and modulation by the cyclic AMP signaling pathway. Cell; **81:** 457-465.

Fan, CM., Lee, CS. and Tessier-Lavigne, M. (1997). A role for WNT proteins in induction of dermomyotome. Dev Biol; **191**: 160-5.

Fanelli, A., Lania, G., Spagnuolo, A. and Di Lauro, R. (2003). Interplay of negative and positive signals controls endoderm-specific expression of the ascidian Cititif1 gene ptomoter. Dev Biol; **263**: 12-23.

Fedi, P., Bafico, A., Nieto Soria, A., Burgess, WH., Miki, T., Bottaro, DP., Kraus, MH. and Aaronson, SA. (1999). Isolation and biochemical characterization of the human Dkk-1 homologue, a novel inhibitor of mammalian Wnt signaling. J Biol Chem; **274**: 19465-19472.

Feldman, B., Gates, MA., Egan, ES. Dougan, St., Rennebeck, G. et al. (1998). Zebrafish organizer development and germ-layer formation require nodal-related signals. Nature; **395:** 181-185.

Finch, PW., Xi, H., Kelley, MJ., Uren, A., Schaudies, RP., Popescu, NC., Rudikoff, S., Aaronson, SA., Varmus, HE. and Rubin, JS. (1997). Purification and molecular cloning of a secreted, Frizzled-related antagonist of Wnt action. Proc Natl Acad Sci USA; **94**: 6770–6775.

Fukushige, T., Hawkins, MG. and McGhee, JD. (1998). The GATA-factor elt-2 is essential for formation of *C. elegans* intestine. Dev Biol; **198**: 286-302.

Gardner, RL. (1982). Investigation of cell lineage and differentiation in the extraembryonic endoderm of the mouse embryo. J Embryol Exp Morphol; **68**: 175 - 198.

Graber, JH., Cantor, CR., Mohr, SC. and Smith, TF. (1999). In silico detection of control signals: mRNA 3'-end-processing sequences in diverse species. Proc Natl Acad SciUSA;**96**:14055-60.

Graff, JM., Thies, RS., Song, JJ., Celeste, AJ. and Melton, DA. (1994). Studies with a *Xenopus* BMP receptor suggest that ventral mesoderm-inducing signals override dorsal signals in vivo. Cell; **79:** 169-179.

Grapin-Botton, A. and Melton, DA. (2000). Endoderm development: From patterning to organogenesis. Trends Genet; **16**: 124-130.

He, X., Saint-Jeannet, JP., Wang, Y., Nathans, J., Dawid, I. and Varmus, H. (1997). A member of the Frizzled protein family mediating axis induction by Wnt-5A. Science; **275**: 1652-1654.

Heasman, J., Wylie, CC., Hausen, P. and Smith, JC. (1984). Fates and states of determination of single vegetal pole blastomeres of *X. laevis*. Cell; **37**:185–194.

Heasman, J. (1997). Patterning the *Xenopus* blastula. Development; **124**: 4179-4191.

Heinemeyer, T., Chen, X., Karas, H., Kel, AE., Kel, OV., Liebich, I., Meinhardt, T., Reuter, I., Schacherer, F. and Wingender, E. (1999). Expanding the TRANSFAC database towards an expert system of regulatory molecular mechanisms. Nucleic Acids Res; **27**: 318-322.

Heisenberg, CP., Tada, M., Rauch, GJ., Saude, L., Concha, ML., Geisler, R., Stemple, DL., Smith, JC. and Wilson, SW. (2000). Silberblick/Wnt11 mediates convergent extension movements during zebrafish gastrulation. Nature; **405**: 76-81.

Henry, GL. and Melton, DA. (1998). Mixer, a homeobox gene required for endoderm development. Science; **281**: 91-96.

Hino, K., Satou, Y., Yagi, K. and Satoh, N. (2003). A genomewide survey of developmentally relevant genes in *Ciona intestinalis*. VI. Genes for Wnt, TGFbeta, Hedgehog and JAK/STAT signaling pathways. Dev Genes Evol; **213**: 264-72.

Hirano, T. and Nishida, H. (2000). Developmental fates of larval tissues after metamorphosis in the ascidian, *Halocynthia roretzi* II. Origin of endodermal tissues of the juvenile. Dev Genes Evol; **210**: 55-63.

Ho, RK. and Kimmel, CB. (1993). Commitment of cell fate in the early zebrafish embryo. Science; **261**: 109-111.

Hoang, B., Moos, M., Vukicevic, S. and Luyten, FP. (1996). Primary structure and tissue distribution of FRZB, a novel protein related to Drosophila frizzled, suggest a role in skeletal morphogenesis. J Biol Chem; **271**: 26131-26137.

Hoang, BH., Thomas, JT., Abdul-Karim, FW., Correia, KM., Conlon, RA., Luyten, FP. and Ballock, RT. (1998). Expression pattern of two Frizzled-related genes, Frzb-1 and Sfrp-1, during mouse embryogenesis suggests a role for modulating action of Wnt family members. Dev Dyn; **212**: 364-72.

Hobert, O. and Westphal, H. (2000). Functions of LIM-homeobox genes. Trends Genet; **16**: 75-83. Review.

Horb, ME. and Thomsen, GH. (1997). A vegetally localized Tbox transcription factor in *Xenopus* eggs specifies mesoderm and endoderm and is essential for embryonic mesoderm formation. Development; **124:** 1689–1698.

Howard, EW., Newman, LA., Oleksyn, DW., Angerer, RC. and Angerer, LM. (2001). SpKrl: a direct target of β-catenin regulation required for endoderm differentiation in sea urchin embryos. Development; **128:** 365-375.

Hu, E., Zhu, Y., Frederickson, T., Barnes, M., Kelsell, D., Beeley, L. and Brooks, D. (1998). Tissue restricted expression of two human Frzbs in preadipocytes and pancreas. Biochem Biophys Res Commun; **247**: 287–293.

Hudson, C., Clements, D., Friday, RV., Stott, D. and Woodland, HR. (1997). Xsox17α and -β mediate endoderm formation in *Xenopus*. Cell; **91:** 397-405.

Hyde, CE. and Old, RW. (2000). Regulation of the early expression of the *Xenopus* nodal-related 1 gene, Xnr1. Development; **127**: 1221-1229.

Iannaccone, P.M., Zhou, X., Khokha, M., Boucher, D., and Kuehn, M.R. (1992). Insertional mutation of a gene involved in growth regulation of the early mouse embryo. Dev Dyn; **194:** 198-208.

Illies, MR., Peeler, MT., Dechtiaruk, A. and Ettensohn, CA. (2002). Cloning and developmental expression of a novel, secreted frizzled-related protein from the sea urchin, *Strongylocentrotus purpuratus*. Mech Dev; **113**: 61-64.

Imai, KS., Takada, N., Satoh, N., Satou, Y. (2000). β-catenin mediates the specification of endoderm cells in ascidian embryos. Development; **127**: 3009-3020.

Imai, KS., Satoh, N. and Satou, Y. (2002). Early embryonic expression of FGF4/6/9 gene and its role in the induction of mesenchyme and notochord in *Ciona savignyi* embryos. Development; **129**: 1729-1738.

Imai, KS. (2003). Isolation and characterization of β-catenin downstream genes in early embryos of the ascidian *Ciona savignyi*. Differentiation; **71**: 346-360

Ishitani, T., Kishida, S., Hyodo-Miura, J., Ueno, N., Yasuda, J., Waterman, M., Shibuya, H., Moon, R. T., Ninomiya-Tsuji, J. and Matsumoto, K. (2003). The TAK1-NLK mitogen-activated protein kinase cascade functions in the Wnt-5a/Ca(2+) pathway to antagonize Wnt/beta-catenin signaling. Mol Cell Biol; **23**: 131-139.

Jaspard, B., Couffinhal, T., Dufourcq, P., Moreau, C. and Duplaa, C. (2000). Expression pattern of mouse sFRP-1 and mWnt-8 gene during heart morphogenesis. Mech Dev; **90**: 263–267.

Johnson, RL., Laufer E., Riddle, RD. and Tabin, C. (1994). Ectopic expression of Sonic hedgehog alters dorsal-ventral patterning of somites. Cell; **79:** 1165-1173.

Jones, CM., Lyons, KM., Lapan, PM., Wright, CVE., and Hogan, BML. (1992). DVR-4 (bone morphogenetic protein-4) as a posterior ventralizing factor in *Xenopus* mesoderm induction. Development; **115:** 639-647.

Jones, SE. and Jomary, C. (2002). Secreted Frizzled-related proteins: searching for relationships and patterns. Bioessays; **24**: 811-820.

Kaestner, KH., Knochel, W. and Martinez, DE. (2000). Unified nomenclature for the winged helix/forkhead transcription factors. Genes Dev; **14**: 142-146.

Kanai, Y., Kanai-Azuma, M., Noce, T., Saido, TC., Shiroishi, T., Hayashi, Y. and Yazaki, K. (1996). Identification of two *Sox17* messenger RNA isoforms, with and without the high mobility group box region, and their differential expression in mouse spermatogenesis. J. Cell. Biol; **133:** 667-681.

Kanai-Azuma, M., Kanay, Y., Gad, JM., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P. and Hayashi, Y. (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. Development; **129**: 2367-2379.

Kessler, D. and Melton, DA. (1994). Vertebrate embryonic induction: mesodermal and neural patterning. Science; **266**: 596-604.

Kiecker, C. and Niehrs, C. (2001). The role of prechordal mesendoderm in neural patterning. Curr Opin Neurobiol; **11**: 27-33.

Kikuchi, Y., Agathon, A. and Alexander, J. (2001). Casanova encodes a novel Sox-related protein necessary and sufficient for early endoderm formation in zebrafish. Genes Dev; **15**: 1493-1505.

Kim, AS., Lowenstein, DH. and Pleasure, SJ. (2001). Wnt receptors and Wnt inhibitors are expressed in gradients in the developing telencephalon. Mech Dev; **103**: 167–172.

Kim, GJ. and Nishida, H. (1999). Suppression of muscle fate by cellular interaction is required for mesenchyme formation during ascidian embryogenesis. Dev Biol; **214**: 9-22.

Kim, GJ., Yamada, A. and Nishida, H. (2000). An FGF signal from endoderm and localized factors in the posterior-vegetal egg cytoplasm pattern the mesodermal tissues in the ascidian. Development; **127**: 2853–2862.

Kimelman, D. and Griffin, KJP. (2000). Vertebrate mesendoderm induction and patterning. Curr Opin Genet Dev; **10**: 350–356.

Kimmel, CB., Warga, RM. and Shilling, TF. (1990). Origin and organization of the zebrafish fate map. Development; **108**: 581-594.

Kobayashi, K. and Nishida, H. (2001). Nuclear plasticity and timing mechanisms of the initiation of alkaline phosphatase expression in cytoplasm-transferred blastomeres of ascidians. Dev Biol; **234**: 510-20.

Kobayashi, K., Sawada, K., Yamamoto, H., Wada, S., Saiga, H. and Nishida, H. (2003). Maternal macho-1 is an intrinsic factor that makes cell response to the same FGF signal differ between mesenchyme and notochord induction in ascidian embryos. Development; **130**: 5179-5190.

Kofron, M., Demel, T., Xanthos, J., Lohr, J., Sun, B., Sive, H., Osada, S., Wright, C., Wylie, C. and Heasman, J. (1999). Mesoderm induction in *Xenopus* is a zygotic event regulated by maternal VegT via TGF-β growth factors. Development; **126**: 5759-5770.

Kowalesky, A. (1866). Entwicklungsgeschichte der einfachen Ascidien. Mem l'Acad St Petersbourg Ser; **7**: 1-19.

Kuhl, M., Sheldahl, LC., Park, M., Miller, J. R. and Moon, RT. (2000). The Wnt/Ca^{2+} pathway: a new vertebrate Wnt signaling pathway takes shape. Trends Genet; **16**: 279-283.

Kuhl, M., Geis, K., Sheldahl, LC., Pukrop, T., Moon, RT. and Wedlich, D. (2001). Antagonistic regulation of convergent extension movements in *Xenopus* by Wnt/b-catenin and Wnt/Ca^{2+} signaling. Mech Dev; **106**: 61-76.

Ladher, RK., Church, VL., Allen, S., Robson L, Abdelfattah, A., et al. (2000). Cloning and expression of the Wnt antagonists Sfrp-2 and Frzb during chick development. Dev Biol; **218**: 183-198.

Lantz, AK. and Kaestner, KH. (2005). Winged-helix transcription factors and pancreatic development Clinical Science; **108**:195-204.

Laverriere, AC., MacNiell, C., Mueller, C., Poelmann, RE., Burch, JBE. and Evans, T. (1994). GATA-4/5/6: a subfamily of three transcription factors transcribed in the developing heart and gut. J Biol Chem; **269**: 23177–23184.

Leimeister, C., Bach, A. and Gessler, M. (1998). Developmental expression patterns of mouse sFRP genes encoding members of the secreted frizzled related protein family. Mech Dev; **75**: 29–42.

Lescher, B., Haenig, B. and Kispert, A. (1998). sFRP-2 is a target of the Wnt-4 signaling pathway in the developing metanephric kidney. Dev Dyn; **213**: 440-451.

Leyns, L., Bouwmeester, T., Kim, SH., Piccolo, S. and de Robertis, EM. (1997). Frzb-1 is a secreted antagonist of Wnt signaling expressed in the Spemann organizer. Cell; **88**: 747-756.

Leong, K., Brunet, L. and Berk, AJ. (1998). Factors responsible for the higher transcriptional activity of extracts of adenovirus-infected cells fractionate with the TATA box transcription factor. Mol Cell Biol; **8**: 1765-74.

Lin, K., Wang, S., Julius, MA., Kitajewski, J., Moos, M., and Luyten, FP. (1997). The cysteine-rich frizzled domain of Frzb-1 is required and sufficient for modulation of Wnt signaling. Proc Natl Acad Sci USA; **94**: 11196-11200.

Lin, R., Hill, RJ. and Priess, JR. (1998). pop-1 and anterior posterior fate decisions in *C. elegans* embryos. Cell; **92**: 229-239.

Liu, C., Li, Y., Semenov, M., Han, C., Baeg, GH., Tan, Y., Zhang, Z., Lin, X. and He, X. (2002). Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. Cell; **108**: 837-847.

Locascio, A., Aniello, F., Amoroso, A., Manzanares, M., Krumlauf, R. and Branno, M. (1999). Patterning the ascidian nervous system: Structure, expression and transgenic analysis of the CiHox3 gene. Development; **126**: 4737–4748.

Logan, CY., Miller, JR., Ferkowicz, MJ. and McClay, DR. (1999). Nuclear β-catenin is required to specify vegetal cell fates in the sea urchin embryo. Development; **126**: 345-357.

Lu, CC., Brennan, J. and Robertson, EJ. (2001). From fertilization to gastrulation: axis formation in the mouse embryo. Curr Opin Genet Dev; **11**: 384-92.

Lustig, KD., Kroll, KL., Sun, EE. and Kirschner, MW. (1996). Expression cloning of a *Xenopus* T-related gene (Xombi) involved in mesodermal patterning and blastopore lip formation. Development; **122:** 4001–4012.

Maduro, MF., Meneghini, MD., Bowerman, B., Broitman-Maduro, G. and Rothman, JH. (2001). Restriction of mesendoderm to a single blastomere by the combined action of SKN-1 and a GSKβ homolog is mediated by MED-1 and -2 in C. elegans. Mol Cell; **7:** 475-485.

Maeno, M., Ong, RC., Suzuki, A., Ueno, N. and Kung, HF. (1994). A truncated bone morphogenetic protein 4 receptor alters the fate of ventral mesoderm to dorsal mesoderm: Roles of animal pole tissue in the development of ventral mesoderm. Proc Natl Acad Sci.; **91:** 10260-10264.

Marcelle, C., Stark, MR. and Bronner-Fraser, M.. (1997). Coordinate action of BMPs, Wnts, Shh and Noggin mediate patterning of the dorsal somite. Development; **124:** 3955-3963.

Mayor, C., Brudno, M., Schwartz, J. R., Poliakov, A., Rubin, EM., Frazer, KA., Pachter, LS. and Dubchak, I. (2000). VISTA: visualizing global DNA sequence alignments of arbitrary length. Bioinformatics; **16**: 1046-1047.

McMahon, JA., Takada, S., Zimmerman LB., Fan, CM., Harland, RM. and McMahon, AP. (1998). Noggin-mediated antagonism of BMP signaling is required for growth and patterning of the neural tube and somite. Genes Dev; **12:** 1438-1451.

Melkonyan, HS., Chang, WC., Shapiro, JP., Mahadevappa, M., Fitzpatrick, PA., Kiefer, MC., Tomei, LD. and Umansky, SR. (1997). SARPs: a family of secreted apoptosis-related proteins. Proc Natl Acad Sci USA; **94**: 13636-13641.

Miller, J. (2001). "The Wnts." Genome Biology; **3**: 1-15.

Molkentin, JD. (2000). The zinc finger-containing transcription factors GATA -4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression. J Biol Chem; **275**: 38949 -52.

Moon, RT., Brown, JD. and Torres, M. (1997). WNTs modulate cell fate and behaviour during vertebrate development. Trends Genet; **13**: 157-162.

Nakatani, Y. and Nishida, H. (1994). Induction of notochord during ascidian embryogenesis. Dev Biol; **166**: 289-299.

Nakatani, Y., Yasuo, H., Satoh, N. and Nishida, H. (1996). Basic fibroblast growth factor induces notochord formation and the expression of As-T, a Brachyury homolog, during ascidian embryogenesis. Development; **122**: 2023-2031.

Nardelli, J., Thiesson, D. and Fujiwara, Y. (1999). Expression and genetic interaction of transcription factors GATA-2 and GATA-3 during development of the mouse central nervous system. Dev Biol; **210**: 305–321.

Neave, B., Rodaway, ARF., Wilson, SW., Patient, RK. and Holder, N. (1995). Expression of zebrafish GATA-3 (gata3) during gastrulation and neurulation suggests a role in the specification of cell fate. Mech Dev; **51**: 169–182.

Nicol, D. and Meinertzhagen, IA. (1988). Development of the central nervous system of the larva of the ascidian, *Ciona intestinalis* L. II. Neural plate morphogenesis and cell lineages during neurulation. Dev Biol; **130**: 737-66.

Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng; **10**: 1-6.

Niehrs, C. (2001). Solving a sticky problem. Nature; **413**: 787–788.

Nishida, H. and Satoh, N. (1983). Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. I. Up to the eight-cell stage. Dev Biol; **99**: 382-94.

Nishida, H. and Satoh, N. (1985). Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. II. The 16- and 32-cell stages. Dev Biol; **110**: 440-54.

Nishida, H. (1987). Cell lineage analysis in ascidian embryos by intracellular injection of a tracer enzyme. III Up to the tissue-restricted stage. Dev Biol; **121**: 526-541.

Nishida, H. (1992a). Developmental potential for tissue differentiation of fully dissociated cells of the ascidian embryo. Rouxs Arch Dev Biol; **201**: 81-87.

Nishida, H. (1993). Localized regions of egg cytoplasm that promote expression of endoderm-specific alkaline phosphatase in embryos of the ascidian *Halocynthia roretzi*. Development; **118**: 1–7.

Nishida, H. (1994). Localization of egg cytoplasm that promotes differentiation to epidermis in embryos of the ascidian *Halocynthia roretzi* Development; **120**: 235-243.

Nishida, H. (2005). Specification of Embryonic Axis and Mosaic Development in Ascidians. (review) Dev Dyn; **233**: 1177-1193.

Nusse, R. and Varmus, HE. (1982). Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. Cell; **31**: 99-109.

Nusse, R., Brown, A., Papkoff, J., Scambler, P., Shackleford, G., McMahon, A., Moon, R. and Varmus, H. (1991). A new nomenclature for int-1 and related genes: the Wnt gene family Cell; **64**: 231-232.

Nusse, R. The Wnt gene Homepage:http://www.stanford.edu/_rnusse/wntwindow.html, update of September 10th, 2001.

Olsen, CL. and Jeffery, WR. (1997). A forkhead gene related to HNF-3beta is required for gastrulation and axis formation in the ascidian embryo. Development; **124**: 3609-3619.

Ortolani, G. (1955). The presumptive territory of the mesoderm in the ascidian germ. Experimentia; **11**: 445-446.

Ortolani, G. (1971). Sul cell-lineage delle ascidie. Boll Zool; **38**: 85–88.

Parks, CL., Banerjee, S. and Spector, DJ. (1988). Organization of the transcriptional control region of the E1b gene of adenovirus type 5. J Virol; **62**: 54–67.

Pearson, WR. and Lipman, DJ. (1988). Improved tools for biological sequence comparison. Proc Natl Acad Sci USA; **85**: 2444-2448.

Pera, EM. and De Robertis, EM. (2000). A direct screen for secreted proteins in *Xenopus* embryos identifies distinct activities for the Wnt antagonists Crescent and Frzb-1. Mech Dev; **96**:183–195.

Pfeffer, PL., De Robertis, EM. and Izpisua-Belmonte, JC. (1997). Crescent, a novel chick gene encoding a Frizzled-like cysteine-rich domain, is expressed in anterior regions during early embryogenesis. Int J Dev Biol; **41**: 449- 458.

Pinson, KI., Brennan, J., Monkley, S., Avery, BJ. and Skarnes, WC. (2000). An LDL-receptor-related protein mediates Wnt signalling in mice. Nature; **407**: 535-538.

Poleev, A., Wendler, F., Fickenscher, H., Zannini, MS., Yaginuma, K., Abbott, C. and Plachov, D. (1995). Distinct functional properties of three human paired-box-protein, PAX8, isoforms generated by alternative splicing in thyroid, kidney and Wilms' tumors. Eur J Biochem; **228**: 899-911.

Rattner, A., Hsieh, JC., Smallwood, PM., Gilbert, DJ., Copeland, NG., Jenkins, NA. and Nathans, J. (1997). A family of secreted proteins contains homology to the cysteine-rich ligand-binding domain of frizzled receptors. Proc Natl Acad Sci USA; **94**: 2859-2863.

Reiter, JF., Alexander, J., Rodaway, A., Yelon, D., Patient, R., Holder, N. and Stainier, DY. (1999). Gata5 is required for the development of the heart and endoderm in zebrafish. Genes and Dev; **13:** 2983-2995.

Reiter, JF., Kikuchi, Y. and Stainier, DY. (2001). Multiple roles for Gata5 in zebrafish endoderm formation. Development; **128:** 125–135.

Renucci, A., Lemarchandel, V. and Rosa, F. (1996). An activated form of type I serine/threonine kinase receptor TARAM-A reveals a specific signalling pathway involved in fish head organiser formation. Development; **122**: 3735-43.

Reuter, R. and Leptin, M. (1994). Interacting functions of snail, twist and Huckebein during the early development of germ layers in Drosophila. Development; **120**: 1137-1150.

Rijsewijk, F., van Deemter, L., Wagenaar, E., Sonnenberg, A. and Nusse, R. (1987). Transfection of the int-1 mammary oncogene in cuboidal RAC mammary cell line results in morphological transformation and tumorigenicity. EMBO J; **6**:127-31.

Ristoratore, F., Spagnuolo, A., Aniello, F., Branno, M., Fabbrini, F. and Di Lauro, R. (1999). Expression and functional analysis of Cititf1, an ascidian NK-2 class gene, suggest its role in endoderm development. Development; **126:** 5149–5159.

Rubinfeld, B., Albert, I., Porfiri, E., Fiol, C., Munemitsu, S. and Polakis, P. (1996). Phosphorylation of Axin, a Wnt Signal Negative Regulator, by Glycogen Synthase Kinase-3β Regulates Its Stability. Science; **272**: 1023-1026.

Ruffins, SW. and Ettensohn, CA. (1996). A fate map of the vegetal plate of the sea urchin (*Lytechinus variegatus)* mesenchyme blastula. Development; **122**: 253-263.

Salic, AN., Kroll, KL., Evans, LM. and Kirschner, MW. (1997). Sizzled: a secreted XWnt8 antagonist expressed in the ventral marginal zone of *Xenopus* embryos. Development; **124**: 4739–4748.

Satoh, N. (1978) Cellular morphology and architecture during early morphogenesis of the ascidian egg: an SEM study. Biol Bull; **155**: 608–614.

Satoh, N. (1994). Developmental biology of ascidians. Cambridge University Press.

Satoh, N. (2003). The ascidian tadpole larva: comparative molecular development and genomics. Nat Rev Genet.; **4**: 285-295.

Satou, Y., Kusakabe, T., Araki, I. and Satoh, N (1995). Timing of initiation of muscle-specific gene expression in the ascidian embryo precedes that of developmental fate restriction in lineage cells. Develop Growth Differ; **37**: 319-327.

Satou, Y. and Satoh, N. (1999). Developmental gene activities in ascidian embryos. Curr Opin Genet Dev; **9**: 542-547.

Satou Y, Imai, KS and Satoh, N. (2001). Early embryonic expression of a LIM-homeobox gene Cs lhx3 is downstream of b-catenin and responsible for the emdoderm differentiation in *Ciona savignyi* embryos. Development; **128**: 3559-3570.

Satou, Y., Takatori, N., Fujiwara, S., Nishikata, T., Saiga, H., Kusakabe, T., Shin-I, T., Kohara, Y. and Satoh, N. (2002). *Ciona intestinalis* cDNA projects: expressed sequence tag analyses and gene expression profiles during embryogenesis. Gene; **287**: 83-96.

Schier, AF., Neuhauss, SC., Helde, KA., Talbot, WS. and Driever, W. (1997). The one eyed pinhead gene functions in mesoderm and endoderm formation in zebrafish and interacts with no tail. Development; **124:** 327-342.

Schlange, T., Andree, B., Arnold, H. and Brand, T. (2000). BMP2 is required for early heart development during a distinct time period. Mech Dev; **91:** 259-270.

Schneider, S., Steinbeisser, H., Warga, RM. and Hausen, P. (1996). β-catenin translocation into nuclei demarcates the dorsalizing centers of frog and fish embryos. Mech Dev; **57**: 191-198.

Schohl, A. and Fagotto, F. (2002). β catenin, MAPK and Smad signalling during early *Xenopus* development. Development; **129**: 37-52.

Schultheiss, TM., Xydas, S. and Lassar, AB. (1995). Induction of avian cardiac myogenesis by anterior endoderm. Development; **121**: 4203-4214.

Shimauchi, Y., Yasuo, H. and Satoh, N. (1997). Autonomy of ascidian fork head/HNF-3 gene expression. Mech Dev; **69**:143–154.

Shirozu, M., Tada, H., Tashiro, K., Nakamura, T., Lopez, ND., Nazarea, M., Hamada, T., Sato, T., Nakano, T. and Honjo, T. (1996). Characterization of novel secreted and membrane proteins isolated by the signal sequence trap method.Genomics; **37**: 273–280.

Shivdasani, RA. (2002). Molecular regulation of Vertebrate early endoderm development. Dev Biol; **249**: 191-203.

Slack, JMW. (1994). Inducing factors in *Xenopus* early embryos. Curr Biol; **4**: 116-126.

Smith, JC. (1989). Mesoderm induction and mesoderm-inducing factors in early amphibian development. Development; **105**: 665-677.

Spagnuolo, A. and Di Lauro, R. (2002). Cititf1 and endoderm differentiation in *Ciona intestinalis*. Gene; **287**:115-119.

Staal, FJ. and Clevers, HC. (2003). Wnt signaling in the thymus. Curr Opin Immunol; **15**: 204-8.

Stainier, DY. (2002). A glimpse into the molecular entrails of endoderm formation. Genes Dev; **16**: 893-907.

Stennard, F., Carnac, G. and Gurdon, JB. (1996). The *Xenopus* T-box gene, Antipodean, encodes a vegetally localised maternal mRNA and can trigger mesoderm formation. Development; **122:** 4179–4188.

Sulston, JE., Schierenberg, E., White, JG. and Thomson, JN. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. Dev Biol; **100:** 64-119.

Suzuki, A., Thies, RS., Yamaji, N., Song, JJ., Wozney, JM., Murakami, K., and Ueno, N. (1994). A truncated bone morphogenetic protein receptor affects dorsal-ventral patterning in the early *Xenopus* embryo. Proc Natl Acad Sci; **91:**10255-10259.

Takahashi, H., Mitani, HY., Satoh, G. and Satoh, N. (1999). Evolutionary alterations of the minimal promoter for notochord-specific Brachyury expression in ascidian embryos. Development; **126**, 3725–3734.

Takahashi, S., Yokota, C., Takano, K., Tanegashima, K., Onuma, Y. et al. (2000). Two novel nodal-related genes initiate early inductive events in *Xenopus* Nieuwkoop center. Development; **127**: 5319-5329.

Tamai, K., Semenov, M., Kato, Y., Spokony, R., Liu, C. et al. (2000). LDL-receptor-related proteins in Wnt signal transduction. Nature; **407**: 530-535.

Terry, K., Magan, H., Baranski, M. and Burrus, LW. (2000). Sfrp-1 and sfrp-2 are expressed in overlapping and distinct domains during chick development. Mech Dev; **97**: 177-182.

Thorpe, CJ., Schlesinger, A., Carter, JC. and Bowerman, B. (1997). Wnt signaling polarizes an early C.elegans blastomere to distinguish endoderm from mesoderm. Cell; **90**: 695-705.

Torres, MA., Yang-Snyder, JA., Purcell, SM., DeMarais, AA., McGrew, LL. and Moon, RT. (1996). Activities of the Wnt-1 class of secreted signaling factors are antagonized by the Wnt-5A class and by a dominant negative cadherin in early Xenopus development. J Cell Biol; **133**: 1123-1137.

Umbhauer, M., Djiane, A., Goisset, C., Penzo-Mendez, A., Riou JF., Boucaut, JC. and Shi, DL. (2000). The C-terminal cytoplasmic Lys-thr-X-X-X-Trp motif in frizzled receptors mediates Wnt/beta-catenin signalling. EMBO J; **19**: 4944-4954.

Uren, A., Reichsman, F., Anest, V., Taylor, WG., Muraiso, K., Bottaro, DP., Cumberledge, S. and Rubin, JS. (2000). Secreted frizzled-related protein-1 binds directly to Wingless and is a biphasic modulator of Wnt signaling. J Biol Chem; **275**: 4374-4382.

von Stein, OD., Thies, WG. and Hofmann, M. (1997). A high throughput screening for rarely transcribed differentially expressed genes. Nucleic Acids Res; **25**: 2598-602.

Wada, H. and Satoh, N. (1994). Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. Proc Natl Acad Sci USA; **91**: 1801-1804.

Wada, S., Katsuyama, Y., Yasugi, S. and Saiga, H. (1995). Spatially and temporally regulated expression of the LIM class homeobox gene Hrlim suggests multiple distinct functions in development of the ascidian, *Halocynthia roretzi*. Mech Dev.; **51**: 115-126.

Wada, H. and Satoh, N. (2001). Patterning the protochordate neural tube. Curr Opin Neurobiol.; **11**: 16-21.

Wang, YS., Macke, JP., Abella, BS., Andreasson, K., Worley, P. et al. (1996). A large family of putative transmembrane receptors homologous to the product of the Drosophila tissue polarity gene frizzled. J Biol Chem; **271**: 4468- 4476.

Wang, W., Wikramanayake, AH., Gonzalez-Rimbau, M., Vlaahou, A., Flytzanis, CN. and Klein, WH (1996). Very early and transient vegetal plate expression of SpKrox1, a Kruppel/Krox gene from *S. Purpuratus*. Mech Dev; **60**: 185-195.

Wang, S., Krinks, M., Lin, K., Luyten, FP. and Moos, M. (1997). Frzb, a secreted protein expressed in the Spemann organizer, binds and inhibits Wnt-8. Cell; **88**: 757-766.

Warga, RM. and Kimmel, CB. (1990). Cell movements during epiboly and gastrulation in zebrafish. Development; **108**: 569-580.

Warga, RM. and Nüsslein-Volhard, C. (1999). Origin and development of the zebrafish endoderm. Development; **126**: 827–838.

Weber, H., Symes, CE., Walmsley, ME., Rodaway, AR. and Patient, RK. (2000). A role for Gata5 in *Xenopus* endoderm specification. Development; **127:** 4345-4360.

Wehrli, M., Dougan, ST., Caldwell, K., O'Keefe, L., Schwartz, S. et al. (2000). Arrow encodes an LDL-receptor-related protein essential for Wingless signalling. Nature; **407**; 527-530.

Weigal, D. and Jackle, H. (1990). The fork head domain: a novel DNA binding motif of eukaryotic transcription factors. Cell; **63**: 455-456.

Weigel, D., Jurgens, G., Kuttner, F., Seiffert, E. and Jackle, H. (1989). The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the Drosophila embryos. Cell; **57**: 645-658.

Whittaker, JR. (1990). Determination of alkaline phosphatase expression in endodermal cell lineages of an ascidian embryo. Biol. Bull; **178**: 222-230.

Winklbauer, R., Medina, A., Swain, RK. and Steinbeisser, H. (2001). Frizzled-7 signalling controls tissue separation during *Xenopus* gastrulation. Nature; **413**: 856-860.

Wylie, CC., Snape, A., Heasman, J. and Smith, JC. (1987). Vegetal pole cells and commitment to form endoderm in *Xenopus laevis*. Dev Biol; **119**: 496–502.

Xanthos, JB., Kofron, M., Wylie, C. and Heasman, J. (2001). Maternal VegT is the initiator of a molecular network specifying endoderm in *Xenopus laevis*. Development; **128:** 167-180.

Yamada, A. and Nishida, H. (1996). Distribution of cytoplasmic determinants in unfertilized eggs of the ascidian *Halocynthia roretzi*. Dev Genes Evol; **206**: 297-304.

Yamamoto, H., Kishida, S., Kishida, M., Ikeda, S., Takada, S. and Kikuchi, A. (1999). Phospholrilation of axin, a Wnt signal negative regulator, by glycogen synthase kinase-3b regulates its stability. J Biol Chem; **274**: 10681-10684.

Yasuo, H. and Satoh, N. (1993). Function of vertebrate T gene Nature; **354**: 582-583.

Yasuo, H. and Satoh, N. (1998). Conservation of the developmental role of brachyury in notochord formation in a urochordate, the ascidian *Halocynthia roretzi*. Dev Biol; **200**: 158-170.

Zdobnov, EM. and Apweiler, R. (2001). InterProScan--an integration platform for the signature-recognition methods in InterPro. Bioinformatics; **17**: 847-8.

Zhang, C., Basta, T., Fawcett, SR. and Klymkowsky, MW. (2005). SOX7 is an immediate early target of VegT and regulates Nodal-related gene expression in *Xenopus*. Dev Biol; **278:** 526-541.

Zhang, J. and King, ML. (1996). *Xenopus* VegT RNA is localized to the vegetal cortex during oogenesis and encodes a novel T-box transcription factor involved in mesodermal patterning. Development; **122:** 4119-4129.

Zhang, J., Huston, DW., King, ML., Payne, C., Wylie, C. and Heasman, J. (1998). The role of maternal VegT in establishing the primary germ layers in *Xenopus* embryos Cell; **94:** 515-524.

Zhu, J., Fukushige, T., McGhee, JD. and Rothman, JH. (1998). Reprogramming of early blastomeres in to endodermal progenitors by a *C.elegans* GATA factor. Genes and Dev; **12**: 3809-3814.

Aknowlegments

I would like to express my gratitude to my director of studies Dr. Antonietta Spagnuolo for giving me the opportunity to work on this project with her fruitful scientific discussion, her advice and her costant support. I would like to thank Prof. S. Piccolo for helping me with his knowledge during ours meeting. I am grateful to Dr Euan Brown who offered me the opportunity of interesting discussion during the Ph.D. course and for critical reading of the manuscript. I thank in situ service headed to Dr. Rita Marino. I should like to thank Dr. Iamunno Gennaro for the preparation of embryos semi-thin sections and the Molecular Biology Service at the Stazione Zoologica "A. Dohrn". Special thanks go to Dr. Laura Zanetti and Dr. Filomena Ristoratore for continuos encouragments and particularly to Dr. Margherita Branno for her help and her warm willingness.

I am indebted with the Stazione Zoologica "A.Dohrn" for giving me the opportunity to attend the courses held by outstanding scientists.